Dissertation submitted to the Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg, Germany for the degree of Doctor of Natural Sciences

Presented by

Master of Science in Pharmacognosy: Tomoko Ishibashi

Born in: Kurume, Fukuoka, JAPAN

Oral-examination: 6th March 2015

Identification and Characterization of Tumor Initiating Cells in Various Mouse Mammary Tumor Models

Referees

Professor Dr. Ana Martin-Villalba (DKFZ, University of Heidelberg)

Professor Donal O'Carroll (EMBL Monterotondo, University of Edinburgh)

Professor Takashi Hiiragi (EMBL Heidelberg, University of Kyoto)

Professor Dr. Joachim Wittbrodt (University of Heidelberg)

Title of the PhD thesis:

Identification and Characterization of Tumor Initiating Cells in Various Mouse Mammary Tumor Models

PhD candidate:

Tomoko Ishibashi

Affiliation:

EMBL Monterotondo, Italy
All the experiments were carried out in the Jechlinger-Sotillo (JS) Lab in EMBL Monterotondo
jointly supervised by Dr. Martin Jechlinger and Dr. Rocio Sotillo
from January 2011 to December 2014.

I dedicate this thesis to my mother, my principal supervisor Dr. Martin Jechlinger, and lastly but not the least, all the lives that were generated and sacrificed in the animal experiments.

Acknowledgements

I would like to express my gratitude to every person that made writing this thesis possible. Among them, I would like to thank especially my parents for giving me continuous support and the opportunity of further education.

I am grateful to the European Molecular Biology Laboratory where I performed the experiments. I would like to acknowledge the EMBL director general Iain Mattaj, the head of EMBL Monterotondo Mouse Biology unit, Philip Avner, and the previous head of unit Nadia Rosenthal. I would like to express my gratitude to the International PhD program and the members of EMBL Graduate Office for their efficient and warm support including Helke Hillebrand, Matija Grgurinovic, Meriam Bezohra and a former member Milanka Stojkovic.

I would like to express my special thanks to my supervisors Martin Jechlinger and Rocio Sotillo for welcoming me to join in the group (J-S Lab) in the end of 2010, and guiding me through in the last full 4 years of my PhD. It has been a real pleasure to work with them; they are always available and giving me honest advises with their different strengths and points of view. I would like to also give thanks to my Thesis Advisory Committee members, Donal O'Carroll, Takashi Hiiragi, and Joachim Wittbrodt for being very approachable and helpful external advisors.

I would like to acknowledge great support from the EMBL Monterotondo staff members, Alexander Young, Daniel Baugh from the Kitchen, as well as Alexandra Moreira, Flavio Zizzo, Silvia Rossetti, Pascale Beudin from our Administration team. I appreciate the immediate and friendly IT support given by Massimiliano Ceccarelli and Olivier Maurel previously.

I would like to thank all the core facilities in our outstation including the Transgenic core facility led by Pedro Moreira and Michela Ascolani, the former head of the Gene Expression core facility Philip Hublitz, and the Microscope core facility managed by Giulia Bolasco. I would like to give my special thanks to Daniel Bilbao from the FACS core facility, for the advice on experimental design, optimization of FACS procedure and teaching me the operation of FACS Aria. I could say that he is one of the mentors in my PhD; for whatever questions I had I could always go and ask him. I would like to also acknowledge Emerald Perlas, head of our Histology core facility as another mentor in campus. He taught me histology from zero and all tricks from his experience. Together with Agnieszka Piszczek, a former member of the facility, he invited me for coffees with smiles and memorable talks. This work would not be

possible without the co-operation of staff in the animal facility, and I would like to thank Maria Kamber for her extremely efficient work and continuous support, and the team of animal facility care takers including Giuseppe Chiaparrelli (Pino), Stefano Tutti, Mateo Gaetani, Andrea Peschiaroli, Violetta Paribeni, and Roberto Voci.

I would like to thank Jens Stolte, former technician in our group, for his positive atmosphere and all the support he gave me including genotyping optimization, qPCR, MMTV-CreET2 construction, as well as organizing parties and road trips together. I thank Joana Passos, another technician and a great bench mate in our group, for the optimization of co-immuno-fluorescence staining on paraffin embedded tissues (SMA, TROMA antibodies). She also taught me some histological techniques and tumor pathology classification. I would like to acknowledge a diligent former colleague in our group, Martina Mantovan for the primary tumor monitoring and producing data on the survival curve of T-O-NEU, T-O-MYC models. I would like to thank all other former lab members and trainee students including Valentina Cappadonna, Martina Calio, Kripa Shobana Ganesh, and Hazel Quinn. Valentina and I enjoyed not only working together but also explored life of Rome. Martina Calio, a dedicating student, often assisted me for the long lasting *in vivo* transplantation work in the internal operation room. Kripa and Hazel were very giving and caring excellent persons who made me very happy in the daily the working environment. They also helped me by proofreading my thesis.

I would also like to thank my other office mates especially Marcos Morgan and Sabin Antony for their kindness and support throughout the past years. In need, they always gave me help and I appreciated this environment very much. I thank my colleague, Ozge Vargel for her kindness and friendship throughout the past 4 years, in fact, since the day of our PhD interview. I would like to thank the kind support of two lab managers, Kerstin Ganter from Lancrin group, and Andreas Hierholzer from Avner group. They also spend some time for the translation of my PhD thesis abstract in German.

I would like to thank current lab members including Marta Garcia Montero Kristina Havas, Ksenija Radic. In particular, I thank Marta for the joint effort on tamoxifen inducible mouse models and characterization of MMTV-CreERT2 and Axin2-CreERT2 lines as well as some histological analysis amongst many other supports she gave me. I appreciate Kristina's critical and constructive advises on the project. I give a special thanks to Ksenija for all her assistance and very supportive and pleasing company. I would like to acknowledge the current and former member of

Sotillo Lab including Maria Giubettini, Yuanyuan Chen (Charles), Cristina Aguirre, Konstantina Rowald, Marion Libouban, and Sharavan Vishaan Venkateswaran for sharing of working space in laboratory and holding joint lab meetings. In particular, Maria and Charles are great office mates on the other side, and I thank them for their advices and a warm company.

I would like to thank Christophe Lancrin for his constructive and kind advices. I thank also Paul Heppenstall and his lab members for sharing the tissue culture room, and especially the kind cooperation of Concepcion Martinez, and Fernanda de Castro Reis, the former and the current lab manager of Heppenstall Group. I would like to thank the great support of Ignasi Claret in FACS core facility between May and June 2014.

I would like to express thanks to the effort of Halldor Stefansson, Cornelius Gross, Jose Viosca and Shane Jeremiah Morley for the continuation of Science & Society program in EMBL Monterotondo outstation. I appreciated kind company of other colleagues across different groups including Alessandro Ciccarelli, Angelo Raggioli, Maria Esteban Masferrer, Tamara Franklin, Andrea Cerase, Manish Kumar, Stefano Gnan, Marco Tosi, Mayya Sundukova, Claudia Carrieri, Christian Much, Ivayla Ivanova, Lina Vasiliauskaite, Rahul Dhandapani and Maya Shvartsman. I would like to thank former students and post doctoral fellows that were in Monterotondo and for passing me the culture and sprit of EMBL, Apar Jain, Hong Wang, Kasper Rasmussen, Nereo Kalebic, Stefano Comazzetto, Guoying Yang (Aiko), Yang Zhan, Rosa Paolicelli, Bjarki Johannesson, Daniela Cornacchia, Luisa Luciani, Manlio Vinciguerra, James Ding and Marianne Hede. I would like to give thanks to many fascinating former trainees including Paola Pisacane, Madhu Elango, Aditya Sankar and Yang Zhang. I would like to also thank the staff from CNR, EMMA and ICGEB including Engin Bojnik, Ozlem Ibrahimoglu and Anna Ferrara. I would like to thank all the members of staff in Portineria, and cleaning service at the campus too, especially in the evenings when I spent long time in campus.

Lastly but most of all, I would like to give thanks for the opportunity of living in the city of Rome, and at the same time working in an international scientific environment. I thank Marita Bartolazzi and her family for welcoming me in their home and inviting me for familial and seasonal occasions. I learned the Roman way of living such as how to dress, how to cook, how to patiently wait for the trains to arrive, how to drink coffee as well as how to relax and enjoying the day-to-day life of the present moment.

Abstract

Breast cancer is not a single disease as it can be classified into different subtypes according to cellular composition, morphology, proliferative index, genetic lesions and therapeutic responses. The molecular and cellular mechanisms underpinning tumor heterogeneity remain a central question in the cancer biology field. To explain the multitude of breast cancer phenotypes, it has been proposed that tumor-initiating cells (TICs) might originate from different cells within the mammary lineage. Further, different oncogenes might elicit distinct phenotypes in a given cell, adding to the complexity of the disease. Single oncogene amplification is observed in human breast tumors. For instance, the human epidermal growth factor receptor type-2 (HER2) proto-oncogene (also called ErbB2, or Neu) is amplified in about 20% of human breast cancers and associated with aggressive phenotypes and poor prognosis. Similarly, 16% of breast cancers have c-Myc gene amplification. Yet, it remains largely disputed which mammary cell type responds to a specific initiating oncogenic mutation. We decided to take a forward approach by testing two contrasting potent oncogenes, Neu (murine form of HER2) or c-Myc, that are frequently overexpressed in breast cancer patients for their capacity to transform different cells of the mammary lineage.

A number of transgenic mouse mammary tumor models have been generated that mimic human breast cancers. Previous studies reported that TICs of Neu driven tumors are the luminal progenitor cells. These studies made use of a mouse model that constitutively overexpressed the Neu oncogene in mouse mammary glands under the MMTV promoter. In contrast, c-Myc overexpressing tumors show a heterogeneous mixed phenotype that is composed of both luminal and myoepithelial cells. Due to the presence of bi-lineage derived cells, we hypothesized that the TICs of c-Myc tumors are the bi-potent stem cells. To clarify the identity and characteristics of TICs, we employed tractable mouse models, Tet-On-Neu/MMTV-rtTA and Tet-On-Myc/MMTV-rtTA, in which we can induce the overexpression of an oncogene at any given time point by administration of doxycycline. The possibility to induce overexpression of an oncogene at adulthood (8-9 weeks after birth) closely recapitulates the timing of somatic mutations acquired by breast cancer patients. In order to demonstrate the direct contribution of distinct breast epithelial cellular lineages to Myc and Neu driven tumorigenesis, we took a FACS approach to separate bi-potent adult stem cells from lineage-committed progenitor cells and differentiated cells. We transplanted the purified cell populations, then induced oncogenes by administration of doxycycline and observed tumor formation *in vivo*.

These tumors obtained from the transplantation experiments were compared with the natural arising primary mammary tumors from Tet-On-Neu/MMTV-rtTA and Tet-On-Myc/MMTV-rtTA by their histological and molecular features. In parallel, we also monitored the clonal growth of these sorted cells *in vitro* after oncogene overexpression using organotypic 3D cell culture assays. Taken together and partially in contrast to our initial hypothesis, Neu induced tumors that are composed of luminal cells can originate not only from the luminal-committed progenitors but also from the bi-potent stem cells. Moreover, luminal committed progenitors were able to give rise to c- induced tumors that are composed of both luminal and myoepithelial lineages. We currently validate our findings by employing additional cell surface antibodies to closer define and sub-fractionate mammary cellular lineages and by an *in vivo* lineage tracing approach.

Our results suggest that Neu and c-Myc TICs can arise from different cellular subtypes in the mammary gland. Notably, progenitor cells seem to be able to establish tumors that consist of basal and luminal cells. These results add to the growing notion that genetic predisposition directs cell fate towards distinct breast cancer phenotypes.

Zusammenfassung

Brustkrebs ist eine komplexe Erkrankung, die hinsichtlich ihrer zellulären Komposition, Morphologie, proliferativem Index, genetischer Disposition und Therapierbarkeit, in verschiedene Subtypen klassifiziert werden kann. Die Erforschung der molekularen und zellulären Mechanismen, welche der Tumorheterogenität zugrunde liegen, ist daher ein Hauptanliegen der Krebsforschung. Eine mögliche Erklärung für die Vielfalt der Brustkrebstypen, vermutet man in den sogenannten tumor-iniziierenden Zellen (TIZ) die aus unterschiedlichen Zellen innerhalb der Brustzelltypen entstehen. Desweiteren können verschiedene Onkogene unterschiedliche Phenotypen in ein und derselben Zelle bewirken, was zur Komplexität der Krankheit beiträgt. In vielen humanen Brusttumoren wird eine erhöhte Expression von nur einem einzelnen Onkogen gefunden. So ist zum Beispiel bei ungefähr 20% der humanen Brustkrebserkrankungen das Proto-Onkogen Epidermaler Wachstumsfaktor Rezeptortyp-2 (HER2, auch ErbB2 oder Neu genannt) hoch reguliert und mit aggressiven Phenotypen und schlechten Prognosen assoziiert. Ein weiteres Onkogen, c-Myc, wird bei 16% der Brustkrebserkrankungen erhöht exprimiert. Dabei ist es nicht klar, welche Zellen des Brustgewebes auf eine onkogene Mutation reagiert und damit zur TIZ wird.

In dieser Arbeit wurden die zwei oben genannten wichtigen Onkogene hinsichtlich ihrer Kapazität untersucht, die verschiedenen Zelltypen der Brustdrüse in TIZ zu transformieren.

Dazu wurde eine Anzahl von transgenen Maus-Brusttumormodelle generiert, die den humanen Brustkrebs nachahmen soll. Frühere Studien berichteten, dass TIZ von *Neu* gesteuerten Tumoren, ursprünglich luminale Vorläuferzellen waren. Im dabei benutzten Mausmodell wurde *Neu* über den *MMTV* Promoter überexprimiert. Im Gegensatz dazu zeigen c-Myc überexprimierende Tumore einen eher heterogenen Phenotyp der aus luminalen und myoepithelialen Zellen besteht. Aufgrund der Anwesenheit von diesen beiden Zelltypen vermuteten wir, dass TIZ aus c-Myc Tumoren bi-potente Stammzellen sein müssten. Um die Herkunft von TIZ zu untersuchen, verwendeten wir ein induzierbares Mausmodell, in welchem wir die Überexpression eines Onkogens zu einem bestimmten Zeitpunkt mit der Verabreichung von Doxycyclin induzieren konnten. Die Möglichkeit der zeitlich steuerbaren Überexpression eines Onkogens im erwachsenen Tier erlaubt damit

eine größtmögliche Annäherung an humane Brustkrebsfälle hinsichtlich des biologischen Alters.

Um den direkten Beitrag von verschiedenen Brustepithelzelltypen zur Entstehung eines Tumors zu demonstrieren, wurden bi-potente Stammzellen aus Vorläufer- und differenzierten Zellen mittels Durchflusszytometrie (FACS) isoliert. Diese Zellpopulationen wurden dann transplantiert und die Onkogenexpression induziert, um die Tumorentstehung *in vivo* zu beobachten. Die von den Transplantationsexperimenten erhaltenen Tumore wurden mit den bereits oben beschriebenen direkt induzierten Tumoren verglichen. Parallel dazu wurden in einem *in vitro* Ansatz isolierte bipotente Stammzellen mittels einer organtypischen 3D Zellkultur nach Induktion auf ihre Eigenschaften hin untersucht.

Zusammenfassend lässt sich nun sagen, dass *Neu*-induzierte Tumore, die hauptsächlich aus luminalen Zellen bestehen, nicht nur aus luminalen Vorläuferzellen entstehen können, sondern auch aus bi-potenten Stammzellen. C-Myc induzierte Tumore, die sowohl aus luminalen als auch aus myoepithelialen Zellen bestehen, können hingegen aus luminalen Vorläuferzellen enstehen. Zur Zeit validieren wir unsere Entdeckungen mit Hilfe von zusätzlichen Zelloberflächen Antikörper um die Brustzelltypen besser zu identifizieren und in Subtypen zu unterteilen.

Unsere Ergebnisse zeigen, dass TIZ aus Neu und c-Myc induzierten Tumoren aus unterschiedlichen Zelltypen in der Brustdrüse entstehen können. Auffallend ist, dass Vorläuferzellen scheinbar in der Lage sind Tumore zu bilden, die aus Lumen- und Basalzellen bestehen. Dieses Ergebnis bestätigt, dass genetische Prädispositionen verschiedene Brustkrebs Phenotypen bewirken.

Thesis Index

1. Introduction

1.1 Overview of breast cancer, a complex disease

Breast cancer is one of the most prevalent cancers worldwide[1]. WHO 2013 Report on Global Health Estimates stated that over 508000 women died of breast cancer in 2011. The GLOBOCAN 2012 (estimated age-standardized incidence and mortality rates) reported that in women, breast cancer is the cancer of highest incidence (25.2%) followed by colorectal (9.2%) and lung (8.8%) cancer[2,3]. Moreover, breast cancer is one of the leading causes of cancer deaths in women (14.7%), closely followed by lung cancer (13.8%)[2,3]. Globally, among the female population, breast cancer incidence rate vary significantly between Western Europe (89.7 per 100000) and Eastern Africa (19.3 per 100000)[4]. Although it is rather rare, breast cancer is also found in male population (1 per 100000)[5].

1.2 Breast cancer heterogeneity

Breast cancer like other cancers encompasses a number of diseases. The different breast cancer types can be identified according to their cellular composition, morphology, proliferative index, genetic lesions and therapeutic responses. Breast cancer can be classified into at least five molecular subtypes (Luminal A, Luminal B, HER2 positive, Claudin-low, and Basal-like) and 18 histological subtypes[6-8]. Clinically, it is crucial to identify the molecular subtypes of breast tumor as the treatment option can differ significantly (Introduction 1.4). One of the major factors that is known to determine the molecular subtypes is the genetic mutations (Introduction 1.3). Depending on the type of oncogenes, the resultant tumors may differ. Genetic instability is also proposed to contribute to tumor heterogeneity[9]. Another possible factor that may contribute to the tumor heterogeneity is a presence of Cancer Stem Cells (CSCs), a fraction of tumor cells that can self-renew, similar to the normal stem cells, and give rise to different lineages of tumor cells resulting in clonal tumor heterogeneity[10]. In addition, depending on the tumor initiating cells (TICs) or the tumor cell-of-origin, the response to the oncogenic cue may differ resulting in tumor heterogeneity (Introduction 1.6).

1.3 Tumor driving genetic mutations

Hanahan and Weinberg[11,12] described several hallmarks of cancer including different biological capabilities acquired by human tumors. Among those, self-sufficiency in growth signals caused by mutations and amplifications of oncogenes are potent drivers of human cancers. As it was mentioned above (Introduction 1.2), one of the known factors of the breast tumor heterogeneity is the nature of the oncogenic mutation and/or loss of tumor suppressor genes. Tables 1. and 2. summarize the main oncogenes overexpressed in human breast cancer and the tumor suppressor genes that are lost or mutated. Among these, HER2 (Human Epithelial Receptor 2) and c-Myc (v-myc avian myelocytomatosis viral oncogene homolog) oncogene amplification are major contributors to human breast cancers (Figure 1.). Interestingly, HER2 and c-Myc oncogenes give rise to distinct and rather contrasting tumor phenotypes as further discussed below (Introduction 1.3.1, 1.3.2). Therefore, we focused our study on these two oncogenes.

Table 1. Major oncogenes up-regulated in breast cancer[13]

Oncogenes	**Function**	**Percentage relevance in breast cancer**
HER-2	Tyrosine kinase receptor, play a critical role in development, activate MAPK, PI3K signaling pathways	Amplification 20–30% [14-16]
c-Myc	Transcription factor, regulate genes involved in proliferation, apoptosis, differentiation	Amplification 1-94%, 15.5% in average [17,18]
Cyclin D1	Cell-cycle mediator, regulate cell cycle progression	Amplification 15-20% [19,20]
Ras	GTPase, regulates signaling pathways involved in proliferation, differentiation, apoptosis, cell adhesion and migration	Amplification <5% [21]
PI3K	Lipid Kinase, regulate cellular growth, differentiation, metabolism, cell survival, proliferation cell cycle, and protein synthesis	Mutation 21–40% [22]
MAPK	Protein kinase, regulate cell proliferation, gene expression differentiation mitosis cell survival, apoptosis	Mutation MAP3K1 in 6% [23]
Akt1	Serine-threonine protein Kinase, play a role in cellular growth, proliferation, survival and angiogenesis	Mutation 1.8–8% [22]

Table 2. Major tumor suppressor genes lost or mutated in breast cancer[13]

Tumor suppressors	**Function**	**Percentage relevance to breast cancer**
p27 (CDKN1B, KIP1)	Inhibit cyclin-dependent protein kinases; arrest cell cycle in G	Mutation 1% [13]
BRCA-1/ BRCA-2	Repair DNA double strand breaks	Mutation 5-10% [24]
CHK2	Cell cycle checkpoint kinase, arrest cell cycle upon DNA damage	Mutation 1% [25] (1100delC variant)
p53	Activate DNA repair, induce cell cycle arrest, initiate apoptosis, cell-cycle checkpoint activation	Mutation 20% [26]
ATM	Serine-threonine protein Kinase, phosphorylates key proteins including p53, CHK2, H2AX	Two fold increase in the risk [27]
PTEN	Lipid/protein Phosphatase, negative regulator of Akt signaling pathway	PTEN Loss 30-40% Mutation PTEN <5% [28]
Rb	Retinoblastoma gene, cell cycle control checkpoint at G1 phase	RB1 loss of heterozygosity (LOH) [29]

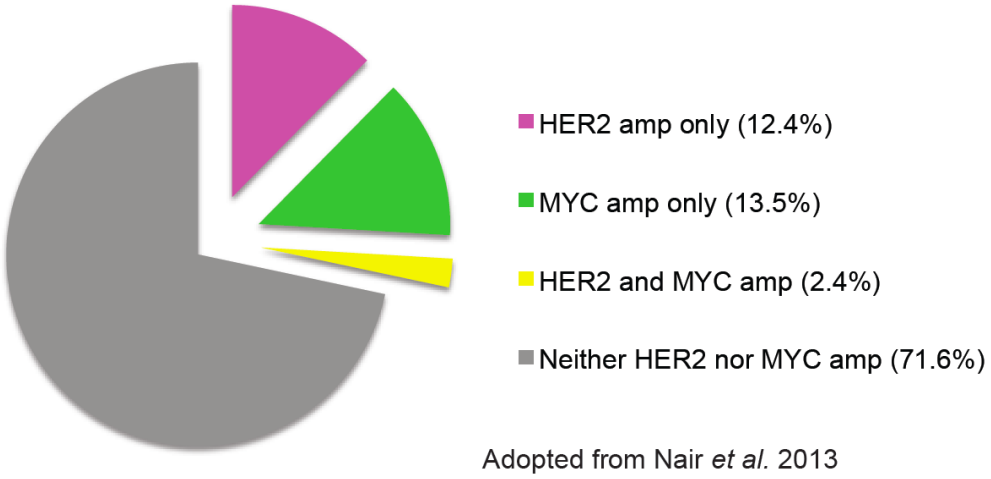


Figure 1. Single- and co-amplification of HER2 and Myc (c-Myc) were present in human breast cancer. Analysis of the Cancer Genome Atlas cohort using the cBio data portal[30].

1.3.1 HER2 oncogene and HER2 driven tumor

As shown in Table 1., HER2 gene amplification is reported in 20-30% of breast cancer cases and correlated with cases of aggressive tumor, and poor prognosis[31,32]. HER2 (also known as Neu, ErbB2) gene is located on chromosome 17q12-q21 and encodes a 185-kDa trans-membrane tyrosine kinase growth factor receptor[33,34]. There are no specific ligands that only bind to HER2 receptor; this results in allowing HER2 to homo- or hetero- dimerize with other members of HER family (HER1, HER3, HER4) in various patterns[35]. Hetero-dimerization HER1-HER2 is common and preferentially activates Mitogen-activated protein kinase (MAPK) pathway that controls both cell proliferation and apoptosis. Another common hetero-dimerization of HER2-HER3 can initiate phosphatidylinositol 3-kinase (PI3K) pathways amongst others, leading to activation of multiple transduction cascades that affect cellular growth, differentiation, metabolism, cell survival, proliferation cell cycle, and protein synthesis[22]. HER2 expression is one of the key criteria for the breast cancer molecular subtype classification. HER2 positive tumors were known to be rather homogeneous, however, recent gene expression studies indicated that almost 50% of HER2 amplified tumors predominantly fall into the clinical Luminal subtypes[36].

1.3.2 Myc oncogene and Myc driven tumors

The c-Myc (Myc) genome amplification found in breast cancer varies; a meta analysis study found its amplification in the range of 1% to 94% (in average

15.5%)[17,18]. In line with the meta analysis, a recent fluorescent in situ hybridization (FISH) study by Nair *et al.* reported that 15% of human breast cancers express Myc gene amplification (Figure 1.)[30]. The percentages of Myc overexpression in breast cancer may increase up to 22-35% at mRNA level, and approximately 40% at protein level[37]. Myc gene is located on chromosome 8q24 and encodes a transcription factor that binds to thousands of promoters. Myc orchestrates various target gene expression that are essential in cellular proliferation, differentiation and apoptosis[38]. Myc overexpressing tumors exhibit high degree of heterogeneity, however, the majority are classified as an aggressive basal-like breast cancer phenotype of with poor prognosis[37,39]. Moreover, Myc overexpression is associated with other oncogenes and tumor suppressor gene mutations. Myc amplification is found in 53% of BRCA1 (breast cancer 1, early onset) mutated basal-like breast tumors[40]. A small fraction (2.4%) of HER2 and Myc co-amplification was are also described (Figure 1.)[30].

1.4 Breast cancer treatment options to date

Since mammary tumor is a heterogeneous disease, a wide variety of treatment options are available depending on the breast tumor subtypes. Current standard breast cancer treatments include surgical removal of the tumor (breast-conserving surgery or mastectomy), radiation therapy, chemotherapy, hormone therapy and targeted therapy. Here are the examples of therapies that are specific for the tumor subtypes. Tamoxifen, a selective estrogen receptor (ER) modulator, and aromatase inhibitors that prevent conversion of testosterone into estrogen are examples of hormone therapy used in breast cancers that are categorized as ER positive[41].

Over the last three decades (from 1980s onwards), it became evident that some tumors are dependent on certain driving oncogenic mutations for their growth and maintenance[42,43]. As described above, in human breast carcinoma, there is a clear involvement of oncogene as the driver of diseases progression. Development of so-called targeted therapies, blocking the specific oncogenic signaling pathway, has improved the treatment of the oncogene addicted breast tumor subtypes significantly.

As one of the examples, in case of HER2 positive tumor subtypes, a humanized recombinant monoclonal antibody, trastuzumab is used and combination therapy of trastuzumab with other chemotherapeutic agents increases survival and therapeutic response rates[44]. The mechanism of action of trastuzumab seems to include many processes such as cytotoxicity effect, inhibition of receptor dimerization, stimulation

of HER2 endocytosis and removal of HER2 from the cell membrane[45,46]. Trastuzumab was shown to bind to the juxtamembrane region of HER2 and block their kinase activation[47]. Pertuzumab, a newer generation of monoclonal antibody was developed owning a different bind sites from trastuzumab on HER2. Pertuzumab prevents the dimerization of HER2 with other HER family members (HER1, 2 and 4)[48]. Targeted therapy is frequently used in conjunction with other conventional therapies including surgery, chemotherapy and radiation therapy. A combination therapy of pertuzumab with trastuzumab and docetaxel significantly improved prolonged progression-free survival of HER2 positive metastatic breast cancer patients with no major cardio-toxicity that was the main adverse affect in the Trastuzumab mono-therapy[49]. A small molecule Lapatinib, a dual kinase inhibitor of HER1 (EGFR) and HER2, is also shown to be effective for the treatment of HER2 positive breast tumors.

Despite of the considerable therapeutic improvement, tumor cells eventually develop a drug resistance not only to the chemotherapeutic agents but also to the targeted therapy. Moreover, the mortality rate of breast cancer remains high due to the issue of tumor recurrence. This could be due to the failure to eradicate all the tumor cells especially the ones that are responsible for tumor initiating and progression. One potential reason for tumor relapse roots in the existence of CSCs, also mentioned above. The CSC hypothesis is an evolving concept in tumor biology. It was first identified in acute myeloid leukemia (ALM) that a fraction of tumor cells are able to self-renew and give rise to different lineage of cancer cells[50,51]. Increasing evidence support that the majority of cancer cells are non-tumorigenic, and only a small fraction of tumor cells are capable of forming new tumors upon transplantation. A number of studies attempt to identify the CSCs that are responsible for relapse of tumor, resistance to the therapy, and metastasis to other organs. In the cases that tumors are dependent on their CSCs for their proliferative capacity, targeting the CSCs have a therapeutic implication. Since current standard therapies target the bulk of tumor cells, development of the CSC specific targeted therapy could be a novel and efficient therapeutic approach.

1.5 Mouse model to study human cancer

Currently, a number of genetically engineered mammary tumor mouse models are available for studying genetic aberrations found in human breast cancer. However, it is essential to identify murine models that conserve clear relevance to human cancer. A recent transcriptomic classification study identified murine models and their human

breast cancer subtype counterparts[7]. Overexpression of oncogenes such as Neu (Murine form of HER2) and Myc in mouse mammary gland can lead to the development of mammary tumors that mimic respective human breast cancer subtypes.

In the last three decades, genetically engineered mouse models (GEMMs) have been used increasingly for better understanding of human cancers[52]. Transgenic mouse models that overexpress oncogenes such as Myc[53] or knock-out mouse models that lack tumor suppressor genes such as p53[54,55] are valuable tools to study the mechanisms of tumorigenesis, tumor recurrence, metastasis, drug resistance, biomarkers, and drug screenings.

Several conventional transgenic mouse mammary tumor models have been described including MMTV-Neu[56,57] and MMTV-c-Myc[58]. In these models, the overexpression of the oncogene is achieved by tissue specific mouse mammary tumor virus (MMTV) promoter in a constitutive manner from the embryonic stage. MMTV promoter is expressed in mammary epithelial cells as well as striated ductal cells of the salivary gland. MMTV promoter has been the most widely used mammary tissue specific promoter, however, there are other mammary tissue specific promoters including whey-acidic-protein (WAP)[59] and Beta-lactoglobulin (BLG)[60] expressed during lactation.

In 1999, tractable transgenic mouse models were first reported (e.g. Myc in hematopoietic cells, $Hras^{G12V}$ in the skin)[61-63]. In contrast to constitutive oncogene overexpression, the inducible transgenic systems allow the induction and de-induction of transgene expression at any given time point. One of the examples is the tetracycline inducible system where the original tetracycline-controlled transcriptional activator (tTA) is a regulator with tight control of target gene expression (Tet-Off system) and its expression is turned off in the presence of doxycycline. In contrast, the reverse tetracycline-controlled transcriptional activator (rtTA) activates the responsive elements only in the presence of doxycycline (Tet-On system)[64]. In case of the Tet-On system, these three following components are needed in order for them to function (Figure 2); 1) the rtTA under the control of the tissue specific promoters such as Mouse Mammary Tumor Virus (MMTV), 2) the Tet-Operator (TetO) repeats cloned upstream the transgene of interest, 3) the tetracycline/doxycycline that allow the rtTA to bind to the TetO repeats and drive transgene expression. In contrast to the constitutive oncogene overexpressing mouse models, this Tet-On system is an attractive tool that, for example, allows us to

study the transition between normal physiological state and tumorigenic state from the adulthood. Moreover, the withdrawal of doxycycline can mimic an oncogene-targeted therapy. Further studies such as tumor oncogene addiction and tumor relapse can be carried out using this GEMM.

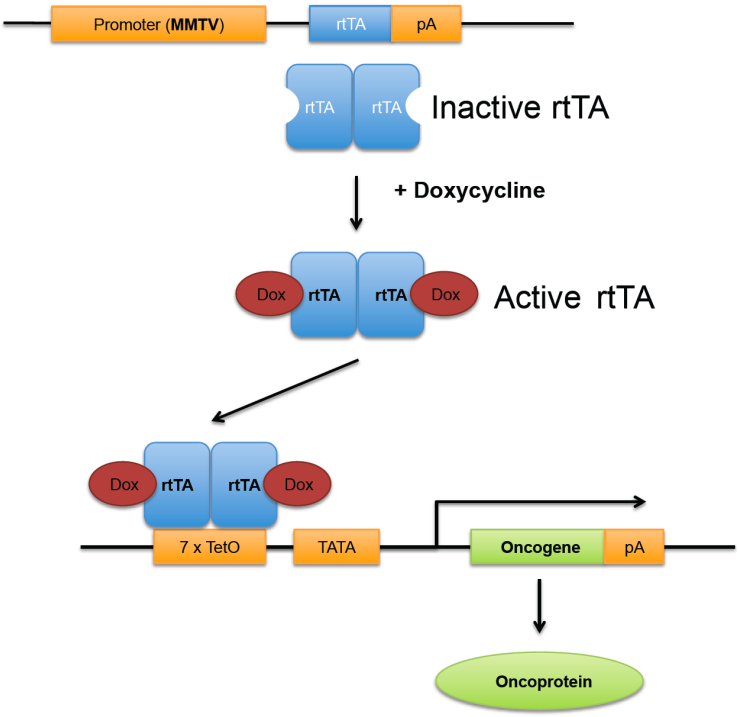


Figure 2. Schematic representation of tetracycline inducible Tet-On system. In the presence of Doxycycline, rtTA (reversed tetracycline trans-activator) binds to TetO repeat that allow the transcription of the target gene (oncogene). For instance, rtTA can be driven by MMTV mammary gland tissue specific promoter.

One of the examples of conventional knock-out mouse model is Rag1$^{-/-}$ immuno-compromised mice. In this model, the recombination-activating genes (Rag) that encode enzymes involved in the process of VDJ recombination are disrupted in a constitutive manner[65]. Therefore, Rag1$^{-/-}$ mice lack B and T lymphocytes. Alternatively, the conditional knock-out models can eliminate expression of a functional gene in tissue specific manner by the use of either loxP site for Cre recombinase or Frt site for Flp recombinase. Unlike the tetracycline inducible system, once the recombination takes place, the excision is irreversible. Yet, conditional knock-outs are frequently used for studying loss-of-function of specific genes. In cancer research, these mouse models revealed the functions of a number of tumor suppressor genes including p53, Rb, BRCA-1, and BRCA-2[66].

The Cre/loxP or Flp/FRT recombinase can also be used to conditionally express a mutant gene of interest, referring to the knock-in models. In contrast to the

conventional transgenics that involves random insertion, a knock-in is a targeted insertion of gene into a specific locus. In addition, the modification of recombinase to contain an estrogen-responsive moiety, Cre-ER or Cre-ERT2, enabled to control the gene activation by the administration of estrogen analogue, tamoxifen[67].

The use of a combination of mouse genetics enabled the current lineage tracing approach possible. Lineage tracing is an identification of progeny that originates from a single cell[68]. A single cell is genetically labeled and the label is transmitted to all its progeny, resulting in a set of labeled clones. In 2007, Clevers and his colleagues identified a Wnt-targeted gene Lgr5 as a marker of the intestinal stem cells generating a knock-in mouse, Lgr5-eGFP-CreERT2[69]. In this model, all the Lgr5 positive cells expressed EGFP (enhanced Green Fluorescent Protein). However, further cross with a Cre-reporter line and generation of a bi-transgenic mouse model Lgr5-eGFP-CreERT2/Rosa26-LacZ allowed the labeling of a small fraction of Lgr5 positive stem cells to express LacZ upon administration of tamoxifen[70]. This approach permits the tracing of the progeny of the Lgr5 positive stem cells during the intestinal development. Furthermore, a combination of this bi-transgenic mouse model with $APC^{flox/flox}$ colon cancer mouse model[71] (loss of function of Adenomatous polyposis coli (APC) gene caused by Cre-loxP mediated recombination) led to the identification of the contribution of Lgr5 positive stem cells in the intestinal tumorigenesis[72].

There are varieties of Cre-reporter mouse models currently available including Rosa26-LacZ, Rosa26-EGFP[73] and Rosa26-YFP[74]. Recently created Rosa26-Brainbow2.1 (Confetti) reporter line would further provide us the opportunity to study the heterogeneity of individual clones by multi-colored fluorescent protein labeling[75,76]. This new generation of mouse genetics is an attractive tool to study not only cell and developmental biological questions but also in contest of tumor biology.

1.6 Tumor initiating cells (Tumor Cell-of-Origin)

Accumulating evidences support that tumors of distinct subtypes could be derived from different “cell-of-origin”[77]. As described above, one of the factors contributing to tumor heterogeneity is the different driving oncogenic mutations[78]. However, one oncogene could give rise to heterogeneous tumors that are classified into different molecular subtypes. To explain the multitude of breast cancer phenotypes, it has been postulated that the differences in the tumor cells-of-origin or the Tumor Initiating Cells (TICs) might play a role.

There are similarities between CSCs and normal stem cell characteristics (self renewal and differentiation capacity), however, CSCs do not necessarily originate from stem cells[79]. CSCs are a fraction of tumor cells that have an ability to self-renewal and differentiate whereas TICs or the tumor cell-of-origin are the normal cells that acquire the first genetic mutation that initiate tumorigenic proliferation[77]. It is a long-standing question in the tumor biology field that which cell types would first respond to somatic oncogenic mutation in early stages of mammary tumorigenesis. Any of the cell types within the breast epithelial hierarchy could serve as the TICs, however, the cell-of-origin may differ depending upon the given genetic mutations and tumor subtypes.

1.6.1 Tumor initiating cells in BRCA-1 mutated tumors

BRCA-1 mutation (185delAG, 5382insC) was first identified as the genetic predisposing mutation for increasing the risk of breast and ovarian cancers[80]. BRCA mutations are found in approximately 5-10% of breast cancer[81]. The TICs or the cell-of-origin of BRCA-1 mutation driven breast tumors is one of the well-studied examples. Although the stem cells or myoepithelial/basal cells are assumed to be the TICs of basal-like tumors, the population that altered their growth properties upon BRCA-1 mutation was, unexpectedly, the luminal progenitor cells[82]. A further study using mouse mammary tumor models bearing inactivation of BRCA-1 with $p53^{+/-}$ background found that the luminal cell population initiated basal-like tumors that resembles the tumors found in BRCA-1 mutation carriers[83]. More recently, using non-plastic human tissue from BRCA-1 mutation carriers, BRCA-1 mutated luminal cells developed basal-like tumors upon transplantation into the humanized mouse mammary fat pad[84]. This study also suggested a possible mechanism of cellular lineage transition by up-regulation of Slug protein expression. All three studies suggest that BRCA-1 mutation can influence cellular differentiation of the luminal progenitor population toward a myoepithelial-basal like state. In addition, it suggests that the tumor phenotypes may not directly reflect the original cellular lineage where TICs reside.

1.6.2 Tumor initiating cells in HER2 driven tumors

Various studies using mouse models attempted to identify the cellular origin of TICs that give rise to HER2 positive tumors. However, it became a controversial question in the field[85,86]. In 2004, partially-identified mammary epithelia cells (PI-MECs) were suggested as the targets of MMTV-NEU driven tumors[87]. PI-MECs are a fraction of

cells that are identified in non-pregnant, parous mice, Wap-Cre/Rosa-LacZ. The identity of PI-MEC was further revealed as a pregnancy-induced stem-like population by the cell surface protein expression, $CD24^{positive}/CD49^{hi}$, identical to the adult mammary stem cells[88]. Another report by Jeselsohn *et al.* indicated that the target of MMTV-NEU is alveolar lineage committed progenitors[89]. This study suggested cycline D1 activity is critical for the self-renewal and differentiation of lobulo-alveolar development progenitors as well as Neu driven tumorigenesis.

More recently, there are three studies indicated that TICs in MMTV-NEU tumors originate from $CD24^{positive}$ luminal progenitor cells[89-91]. The expansion of $CD24^{high}CD49f^{mid/low}$ luminal progenitor population was found in pre-neoplastic MMTV-HER2/Neu mammary glands and displayed additional luminal lineage specific cell surface marker $Ep\text{-}CAM^{high\ 91}$. Moreover, the gene expression profiles of MMTV-HER2/Neu tumor cells were closely correlated to the gene signature of luminal progenitors.

To date, the majority of studies agree that the luminal progenitor population is the TICs of Neu overexpressing tumors. However, the origin and identity of TICs are yet to be determined, since one of the withdrawals in all those reports is the use of the constitutive Neu overexpressing mouse mammary tumor model. Use of tractable oncogenic mouse models that can closely mimic the somatic mutation of HER2 in patients may clarify the true cell-of-origin of Neu driven tumors.

1.6.3 Tumor initiating cells in Myc driven tumors

Myc amplified human breast tumors are known to exhibit a basal-like phenotype with a high degree of heterogeneity[37,39]. The tumors obtained from MMTV-Myc, a constitutive Myc overexpressing mouse mammary tumor model, also showed a clear heterogeneity confirmed by both histological and gene expression profile[92]. Using an inducible transgenic mouse model (Tet-On-Myc/MMTV-rtTA), Myc driven breast cancer was subjected to the oncogene withdrawal that mimics targeted therapy[93]. Myc-dependent tumor regressed as expected, however, a subset of tumors that acquired *de novo* mutation grew Myc-independent tumors. This would also mimic the cases of acquired drug resistance seen in patients. Myc-independent tumors exhibited epithelial–mesenchymal transition (EMT) phenotype with a metastatic property. This study showed that Myc overexpression could lead to tumor heterogeneity, oncogene independence and acquisition of additional mutations including Ras, TGFβ (Transforming growth factor beta), and TNFα (Tumor necrosis

factor alpha) pathways. However, the origin of TIC of Myc driven tumor is yet to be identified. The composition of Myc primary tumor is a mixed basal phenotype, the mammary stem cells or their myoepithelial lineage are the closest candidate.

1.7 Mammary gland and their Cellular Hierarchy

1.7.1 Composition of mammary glands

To understand the malignant state, a detailed knowledge of the healthy tissue is necessary. The healthy adult mammary gland is an organized branched system separated by an extracellular matrix (ECM) and basement membrane, and embedded in stroma composed of fibroblasts, adipocytes and other non-epithelial cells. The mammary epithelial tree is composed of two lineages, the 1) myoepithelial/basal cells in the basal layer tree, 2) luminal cells in the inner layer that are sub-divided into excretory luminal cells in ducts and secretory alveoli cells in lobules (Figure 3). Major mammary gland development takes place at following three distinct periods, 1) embryonic stage, 2) puberty, and 3) pregnancy and lactation[94]. In mice, the formation of mammary bud starts from ectoderm approximately by embryonic day (E) 10.5, and the rudimentary ductal tree arises from epidermis in between E13.5 to birth. Secondly, the extensive branching and elongation of the ducts grow via terminal end buds (TEBs) in the fat pad starting 3 weeks after birth throughout puberty until 5-6 weeks after birth. Lastly, the lobulo-alveolar development takes place during late pregnancy and lactation[94].

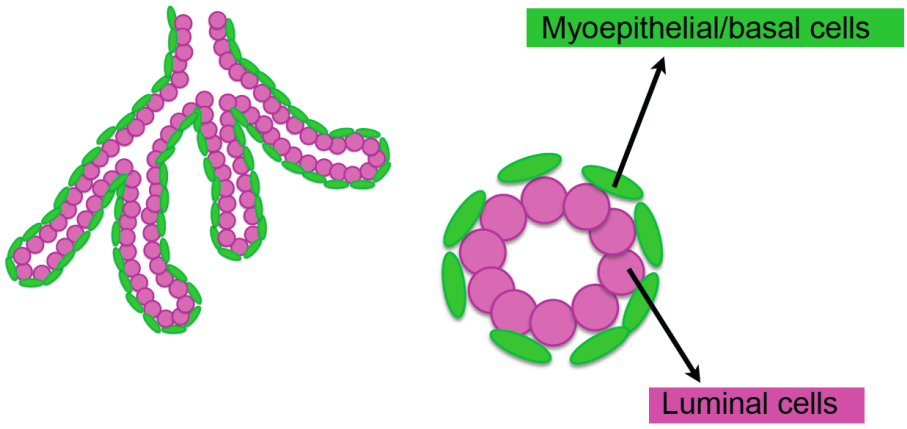


Figure 3. Scheme summarizing the two distinct lineages co-exist in the mammary gland of virgin adult mice. Mammary epithelium is organized as a bilayer ductal system with an inner layer of luminal secretory cells (K8/18 positive in magenta) and an outer layer of basal myoepithelial cells (SMA or K14 positive in green).

During the embryonic stage and puberty, mammary glands rapidly expand and there are significant increases in number of stem cells that are actively proliferating[95,96]. In

addition, mammary glands go through cyclical proliferation and involution during adulthood from puberty till menopause[97]. In mice estrus cycles occur within a 5-day cycle that is equivalent in humans' a 28-day menstrual cycle. There are four phases of estrus cycles including proestrus, estrus, metestrus, and diestrus. Joshi *et al*.[98] reported that each estrus cycle expands the total number of mammary epithelial cells (1.9-fold in diestrus phase). In addition, the number of colony forming unit (putative stem cells) increases in the diestrus phase (14-fold) at the highest and diminish in estrus phase[98].

1.7.2 Mammary cellular lineages

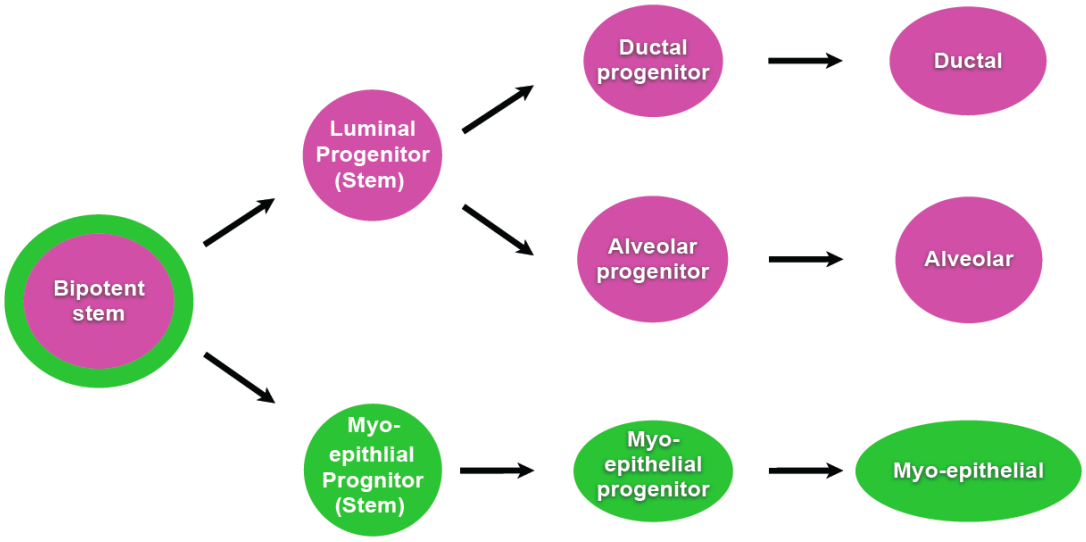
There are specific markers that define two distinct mammary cellular lineages, the luminal (also referred as epithelial) and myoepithelial (also referred as basal). These two different cell types can be distinguished from each other by their cytokeratin expression due to their differentiation status[94]. Keratins are intermediate filament–forming proteins and exist as heteropolymers formed by the non-covalent coiled-coil interaction between one of type I acidic (K9-K28) keratin family and one of type II basic (K1–K8 and K71–K74) keratin family[99,100]. Keratins not only have the structural role to provide cells their shape and rigidity but also regulatory role in cell cycle, apoptosis, cellular stress response, cell size, protein synthesis and membrane trafficking[101-104]. In mammary epithelia, the luminal cells that are responsible for milk production and secretion express cytokeratin of simple (non-stratified) epithelia K8 and K18. Other cytokeratins such as K7 and K19 are also found in luminal cells in the mammary glands[105].

Myoepithelial cells build the basal cell layer of mammary ducts bordering with the basement lamina. Their cytoplasm contains bundles of smooth muscle actin (α-SMA or referred as SMA below) micro-filaments and myosin filaments, which serve as the major contractile apparatus for milk ejection[106]. In addition, accumulating evidence support a variety of functional roles of myoepithelial cells in mammary gland growth regulation, development, and differentiation as well as the control of tumorigenesis[107]. In contrast to the smooth muscle cells, the mammary myoepithelial/basal cells also express the stratified cytokeratin such as K5 and K14[108]. In addition, there are other markers employed to identify myoepithelial lineage including P-cadherin, a cell adhesive molecular[109,110], and p63, a key regulator of cellular adhesion and survival in basal cells[111,112]. Among those myoepithelial markers, the SMA expression in the mammary epithelia is known as fully differentiated myoepithelial marker, whereas K14 is not only expressed in myoepithelial lineages but also among the un-

differentiated luminal cells (i.e. stem/progenitor cells)[113]. Lastly, a member of transcription repressor Snail family, Slug (Snail2) is exclusively expressed in mammary myoepithelial/basal lineage and reported to play an important role in cellular differentiation and mammary gland morphogenesis[114].

1.7.3 Mammary cellular hierarchy

The concept and existence of stem cells in adult tissue were first proposed over 50 years ago in hematopoietic system[115,116]. Similarly, in adult mouse mammary epithelia, the existence of stem cells was proved by transplantation of the tissue fragments[117] and later by serial transplantation of a single cell and in limiting-dilution assay[118] (introduction 1.9.1). In the adult virgin mammary epithelial hierarchy, various distinct mammary epithelial cellular subsets have been described according to their morphology and functionality (Figure 4). Although mammary epithelial cellular hierarchy is an ongoing debate, several studies agree that multi-potent stem cells are present on the top of the hierarchical lineage. However, it is important to note that the stem cell population is not comprised of homogeneous population, but including different sub-class of stem cells (slow-cycling, long-term and proliferating, short-term repopulating cells both of which are multi-potent)[119]. A recent lineage tracing study proposed the existence of uni-potent (luminal or myoepithelial/basal) stem cells that retain their self-renewal capacity and differentiation in its lineage[120]. However, the presence of multi-potent adult mammary stem cells versus lineage-committed stem cells is still highly debated in the field[78,120,121]. Lower in its hierarchy, there are transient amplifying cells, luminal progenitor cells that are restricted to either ductal or alveolar differentiation, and myoepithelial progenitors. At the bottom of the hierarchal tree are further differentiated functional cells including ductal cells, alveolar cells and myoepithelial cells.



Adopted from Visvader & Stingl 2014

Figure 4. Scheme summarizing the hypothetical model of the mammary epithelial hierarchy.

1.8 Cell surface markers

Various cell surface proteins were studied in order to identify specific cellular lineages of the mammary gland. Mouse mammary gland share many features that are similar to human including cell surface protein expressions. We introduce below the major cell surface proteins that are used in the study of mouse mammary gland cellular hierarchy, however, their human counterparts are reviewed extensively elsewhere[119,122,123].

1.8.1 Stem / Myoepithelial lineage markers

Some of the classical methods used to isolate stem cells are label retention with BrdU labeling or other label retaining dye, stem cell antigen Sca-1 expression, or isolation of the "side population" that can efficiently exclude Hoechst 33342 or rhodamine 123 dyes[124]. Lipophilic dye such as PKH have been reported to enrich putative stem cells by their label retaining characteristics[125,126]. Increased aldehyde dehydrogenase (ALDH) activity is known as one of the characteristics of stem/progenitor properties[127].

Much effort has been made to identify various cell surface proteins that can distinguish different mammary epithelial cell population[119] (Figure 5). In 2006, a combination cell surface markers was described to enrich mouse mammary stem cells for the first time[118,128]. Purified stem cells in the mammary epithelia were

characterized by the combined expression of CD29^{hi} (Integrin beta 1), CD24^{mid} (heat stable antigen), and CD49f^{hi} (Integrin alpha 6) presenting the estimated mammary repopulating unit (MRU) of 1/60-90[118,128] upon transplantation.

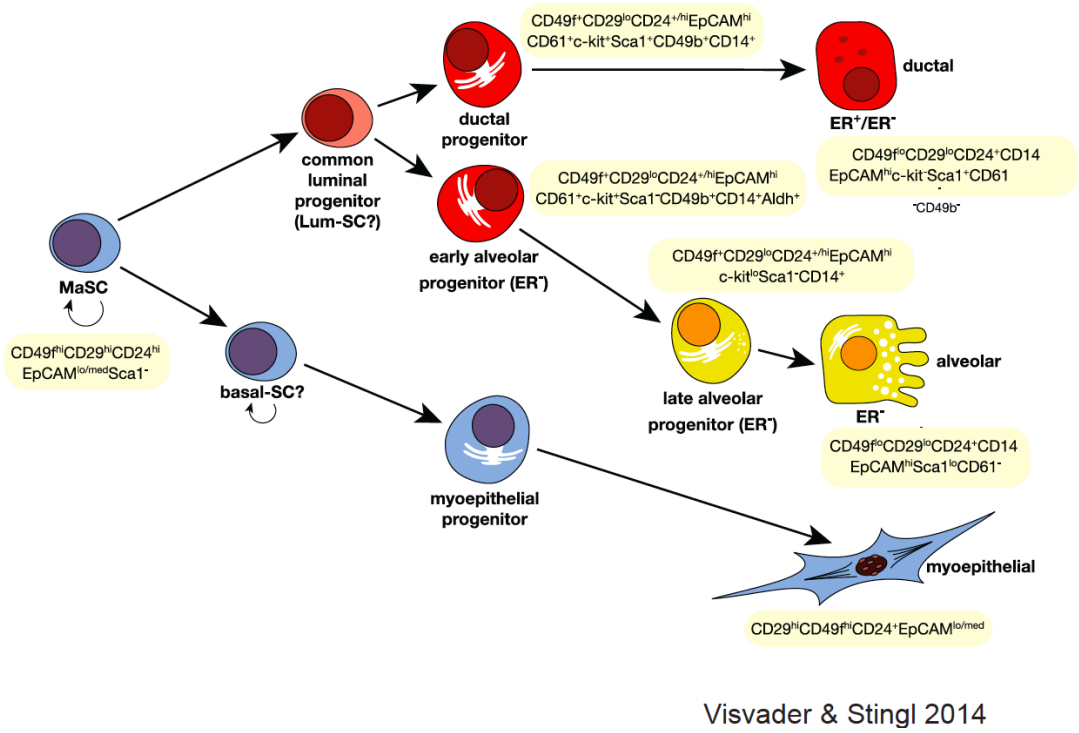


Figure 5. Scheme summarizing the cell surface markers used for the isolation of currently identified epithelial cell subsets from the mouse mammary gland. (ER = Estrogen receptor alpha)

Multiple studies agree that the stem cell reside within the myoepithelial lineage[129-131]. However, there are currently no specific cell surface markers to delineate myoepithelial progenitors or differentiated myoepithelial cells. Gene expression analysis of stem cell population showed no significant difference from their myoepithelial population[128], however, recent RNA sequencing of the stem cell population combined with a long-term H2B label-retaining assay identified one cell surface protein such as CD1d, claimed as the marker for further purified mammary bi-potent stem cells[132]. CD1d is a glycoprotein expressed in antigen-presenting cells involved in T-cell antigen presentation. The refined CD1d positive stem cell population was shown to be able to reconstitute the MG in the efficiency of 1/44.

1.8.2 Luminal lineage markers

Increased expression of CD24 is a widely accepted luminal cell marker together with EpCAM[82,129]. Other cell surface markers that are reported to distinguish luminal progenitor cells from differentiated ductal and alveolar cells are CD61[132,133], CD133[132], CD49b[134], CD14[134,135], c-Kit[135,136] and Sca1[134,137]. Estrogen receptor alpha is expressed among the luminal lineage but not in the alveolar[119].

1.9 Approaches to test cells of the mammary lineage for their stem cell and progenitor cell capacity

1.9.1 Transplantation assay

Mammary gland is a unique organ that allows transplantation and complete reconstitution of mammary gland using cleared fat pad transplantation assay[117]. The MRU described above is an estimate of mammary gland reconstitution efficiency generated from the results of transplantation and analysis by ELDA (Extreme Limiting Dilution Assay)[118,138]. In addition to the MRU, the full-ness of the mammary fat pad by the repopulated gland can also be assessed for the reconstitution efficiency of the transplanted cells. Over the last half century, the transplantation of cells into cleared mammary fat pads has been the gold standard assay to assess the stem cell self-renewal and differentiation capacity[117,118,128,139]. The endogenous mammary anlage was removed from three weeks old syngeneic or immuno-compromised mice, and the rest of the fat pad was used as the site of transplantation (Figure 6).

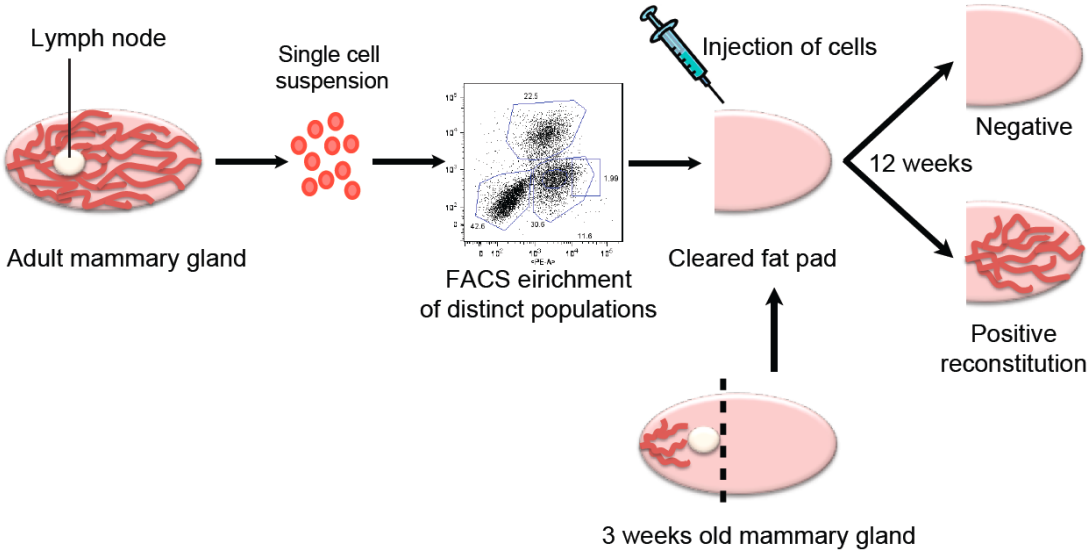


Figure 6. Schematic representation of the *in vivo* transplantation assay for detected adult MRUs. Mammary epithelial cells are dissociated into a single cell suspension, and after FACS enrichment, transplanted into the cleared fat pad of 21 days old Rag1$^{-/-}$ immuno-compromised female mice. 12 weeks later, the glands are removed and scored for the positive reconstitution and fullness of the fat pad with the newly formed mammary glands or negative when the glands are absent.

The transplantation assay was also used in order to assess the cellular tumorigenic capacity in breast CSC research extensively. This orthotopical *in vivo* transplantation site provides much closer microenvironment to the physiological mammary epithelial cells or breast tumor cells in comparison to subcutaneous transplantation[117]. As one of our main read-out assay, the tumor-initiating capacity of normal mammary epithelial sub-populations that experience oncogenic overexpression can be also examined in this procedure (Figure 7.). This would provide valuable information including the origin of TICs, their capacity for self-renew and repopulate tumors. The tumors obtained from transplantation and their primary tumor counterparts can be cross-examined for their phenotypes.

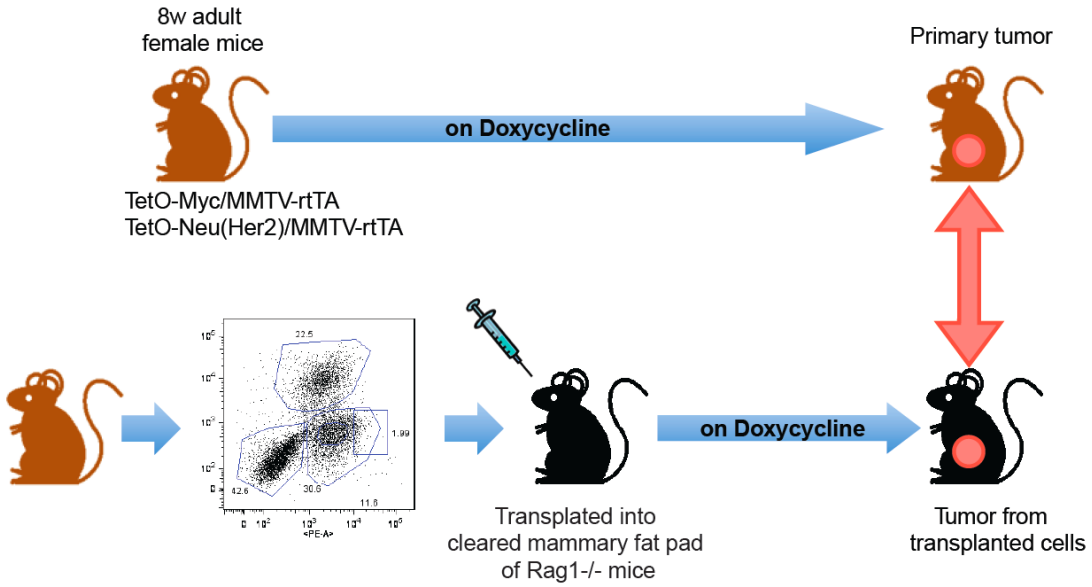


Figure 7. Experimental scheme showing the comparison of FACS isolated mammary epithelial cells giving rise to tumors in immuno-deficient mice and the transgenic mouse models giving rise to primary tumors after administration of doxycycline.

1.9.2 Three-dimensional *in vitro* culture system

As the genetically engineered mouse models permit us to study the effect of oncogene overexpression in tractable manner, we are coming very close to mimic how the events of tumorigenesis and progression occur in patients. However, the

studies conducted with the mouse models have their own limitations for the time that takes a tumor to develop, the cost, as well as the inadequate technology to monitor the mechanistic insights of the cell of interest in detail over a period of time. This has promoted the development of *in vitro* culture system that can closely recapitulate an *in vivo* cellular organization and the microenvironment. Previously, Bissell and her colleagues developed a three-dimensional (3D) culture system of mammary cells that can grow and expand in an organotypic manner maintaining the correct epithelial polarity and other crucial features that are needed for the control of cell proliferation, survival and differentiation[140]. More recently, the 3D culture system was optimized for monitoring the process of induction and de-induction of oncogene in the primary mammary cells from the tractable mouse mammary tumor models[141]. We have used this culture system in order to observe the response of specific cell types before and after the overexpression of oncogenes (Figure 8.).

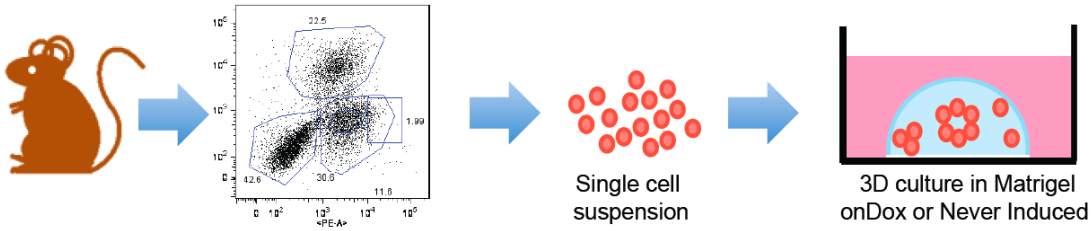


Figure 8. Scheme of the experimental procedure of *in vitro* 3D culture of FACS sorted cells.

1.9.3 Genetic stem cell markers

Although the *in vivo* transplantation has been widely used as a gold standard assay in previous studies, a recent lineage tracing study suggested that the transplantation experiments does not necessarily reflect physiological behavior and cell-fate and artificially influence the ratio of differentiation and regenerative potential[120]. As the result of this transplantation "artifact", lineage-tracing approach are increasingly employed in order to identify all progeny of a single cell such as stem cells and assess their nature and function in a physiological context. This approach allows us to study the labeled specific cell population and follow them overtime upon overexpression of oncogene.

There are several potential genetic stem cell markers of the mammary glands. Lrp5 is expressed in the fraction of stem cells and maintain their myoepithelial lineage[142]. Rohrschneider and his colleague identified s-SHIP promoter that mark actively

proliferative stem cells in the terminal end buds of mammary alveolar cap cells[143]. More recently, the lineage tracing approach was used employing two distinct lineage specific markers, K8 (luminal) and K14 (myoepithelial/basal) by the Blanpain group[144]. This tracing approach of luminal and myoepithelial lineage resulted in identification of uni-potent stem cells in the adult mammary gland. It has been rather a disappointment in the field that Lgr5, an established genetic marker of stem cells in the intestine and skin, was not specifically expressed in the stem cell population[130,144-146]. Other Wnt target genes such as Axin2[146] and Protein C receptor[130] are the two of the currently established bi-potent stem cell makers in the mammary gland. Interestingly, at the adult stage (8 weeks old) Axin2 is expressed among the stem cells, however this differs depending on the developmental stage. At E14, Axin2 will mark luminal cell population, and at puberty, it is expressed among the myoepithelial population. The limitation of lineage tracing approach is the availability of genetic markers and the construction of mouse models that enable to label the distinct cell population. Nevertheless, this approach would provide the most physiological environment to study a specific mammary epithelial subpopulation and their contribution to breast tumorigenesis.

1.10 Un-known facts, questions to be answered

In this research we aim to answer which cell types in the adult mammary glands are the TICs or tumor cell-of-origin in oncogene (HER2 and Myc) driven breast tumors. In the case of HER2 tumor, the majority of studies suggest that the tumor cell-of-origin is so far proposed as the luminal progenitor cells. However all these studies were conducted using a mouse model that overexpressed the oncogene in a constitutive manner. In order to mimic patients who acquire somatic mutation of HER2 in adulthood, we have employed a tractable mouse mammary tumor models instead. This system enables us to examine how non-plastic mammary epithelial cells transform into tumorigenic cells in respond to overexpression of oncogenes.

In contrast to HER2 driven tumors that are composed of luminal cells, Myc tumors exhibit basal-like phenotype. The TICs of Myc induced tumors are not extensively studied, and we would like to identify their tumor cell-of-origin and if it differs from the HER2 case. In addition, it is not yet clear whether the cell-of-origin determines different tumor subtypes, or regardless from the cell-of-origin, the genetic mutation would determine the tumor subtype.

2. Materials and Methods

2.1 Mouse Models

Genetically engineered and mutant mice strains used were Tet-On-Myc/MMTV-rtTA (T-O-MYC)[147] and Tet-On-Neu/MMTV-rtTA (T-O-NEU)[148], CAG-H_2B-GFP[149], Axin2-CreERT2[131], Rosa26-Brainbow2.1[150], Rosa26-YFP[74], Rag1$^{-/-}$[151] .

2.1.1 Mouse Strains, animal husbandry, genotyping

Mice were generated in mixed background (between C57Bl/6N, 129S2/SvHsd, and FVB/NHanTM Hsd) and tumors were induced through the activation of transgene expression. All the mice were bred and maintained at EMBL Mouse Biology Unit, Monterotondo, RM, Italy following to the current Italian legislation (Article 9, 27. Jan 1992, number 116) under license from the Italian Health Ministry.

Table 3. List of primers used for genotyping

TetO-MYC	myc 5'-3'	TAGTGAACCGTCAGATCGCCTG
	mcy 3'-5'	TTTGATGAAGGTCTCGTCGTCC
TETO-Neu	TAN-IRES 3528F	GACTCTCTCTCCTGCGAAGAATGG
	TAN-IRES 3914B	CCTCACATTGCCAAAAGACGG
MMTV-rtTA	CMV-rtTA F	GTGAAGTGGGTCCGCGTACAG
	CMV-rtTA R	GTACTCGTCAATTCCAAGGGCATCG
EGFP	H2B EGFP (F)	CAAGGGCGAGGAGCTGTT
	H2B EGFP (R)	AAGTCGTGCTGCTTCATGTG
LGR5-GFP K1	LGR5-GFP K1 Fwd	CACTGCATTCTAGTTGTGG
	LGR5-GFP K1 Rev	CGGTGCCCGCAGCGAG
Rag1	Rag1 COMUN	CCGGACAAGTTTTTCATCGT
	Rag1 WT	GAGGTTCCGCTACGACTCTG
	Rag1 MUTANT	TGGATGTGGAATGTGTGCGAG
Axin2	WT FW	AAGCTGCGTCGGATACTTGAGA
	WT RV	AGTCCATCTTCATTCCGCCTAGC
	Cre RV	GCACGTTCACCGGCATCAAC
LacZ	LACZ 1	AAAGTCGCTCTGAGTTGTTAT
	LACZ 2	GCGAAGAGTTTGTCCTCAACC
	LACZ 3	GGAGCGGGAGAAATGGATATG
Brainbow2.1	11341	GAATTAATTCCGGTATAACTTCG
	oIMR8545	AAAGTCGCTCTGAGTTGTTAT
	oIMR08916	CCAGATGACTACCTATCCTC
MMTV-Cre	CreER FW	GCTGGCCCAAATGTTGCTGC
	MMTV RV	TTTGAGTAAACTTGCAACAG

2.1.2 Construction of transgenic lines

MMTV-CreERT2 was generated together with Jens Stolte, Martin Jechlinger, Philip Hublitz from the gene expression core facility and Pedro Moreira from the transgenic core facility in EMBL Monterotondo. MMTV-LTR sequence was taken from Donehower *et al.*[152] and plasmid vector was cloned by Jens Stolte, then injected to mouse embryonic stem cells by Pedro Moreira.

2.1.3 Transgene activation in tractable oncogenic mouse models

Induction of transgene expression was achieved by administration of a doxycycline-supplemented diet (625 mg/kg doxycycline) (Harlan Laboratories) after the female mouse reaches their adulthood (8 weeks). Low dose doxycycline administration was achieved by 0.012mg/ml in drinking water. Upon formation of primary mammary tumor in T-O-MYC and T-O-NEU, mice were sacrifice at the human end-point (1cm mammary tumor development).

2.1.4 Transgene activation by Tamoxifen Induction

Tamoxifen (Sigma, T5648) was dissolved in sunflower seed oil (10mg/ml) (Sigma, #S5007) and freshly prepared each time on the same day of the injection. In order to dissolve the crystal of tamoxifen, the solution was mixed at 42C up to 30 minutes then slowly filtrated with 0.45μm filter. Intraperitoneal Injection of tamoxifen solution was performed on Axin2-CreERT2/Rosa26-LSL-Brainbow2.1, Axin2-CreERT2/Rosa26-YFP bi-transgenic female mice and their experimental controls.

Table 4. Various Tamoxifen dosages used in previous studies.

Different dosage of tamoxifen	Administration method
(High dose)	Three injections of 5 mg every other day, totaling 15 mg (adult)[120].
(Medium dose)	4 mg/25g body weight (adult)[131]. 1 mg for prepubescent mice (injected between P14 and P16), 2 mg for pubescent mice (injected between P28 and P35), 4 mg for adult virgins (injected between P56 and P63). 0.5 mg/25 g body weight single injection of TM to pregnant mothers between E12.5 and E17.5
(Low dose)	Single injection of tamoxifen 1.5 mg (adult)[153].

2.2 Preparation of mammary epithelial cell suspension

Mammary glands were dissected from 7-10 weeks old female virgin mice that were in the diestrus phase of estrus cycle. Preparation of mammary cells to single cell suspension method was modified from Jechlinger *et al.*[154]. In short, the tissue from five mammary glands without mechanical dissociation was placed in 5ml digestion medium (DMEM F-12 medium (Lonza/Amaxa Cat#BE12-719F) with added supplements 25mM HEPES (Gibco Life technologies #15630-056), 5ml Penicillin/Streptomycin (Gibco Life technologies #15140-122), that has 150U/mL Collagenase Type 3 (Worthington Cat#LS004183) and 20mg/ml Liberase TM (Roche Cat#05401127001). Mammary glands were digested overnight for up to 18 hours at 37°C in a 5%(vol/vol) CO_2 atmosphere in loosely capped polypropylene 50ml conical tubes. Mammary glands were re-suspended with mechanical force and diluted with

warm phosphate buffered saline (DPBS). Re-suspension was centrifuged at 300g (around 1200rpm) for 5 minutes at room temperature. Interphase between fat and cell pellet was removed. 15ml of 0.25% trypsin EDTA (Invitrogen Cat#25200-056) was added to the cell pellet re-suspended and incubated for 20 minutes at 37°C in a 5%(vol/vol) CO_2 atmosphere loosely capped. The suspension was mechanically dissociated every 10 min. In order to stop trypsin digestion, cells were washed with 35ml of trypsin deactivation media (DMEM/F12 with L-glutamine, 15mM HEPES, supplemented with 1M HEPES to 25mM final concentration and 10% Tet-Free Serum (Clontech Cat#631106). The solution was treated with 5–15mg/ml DNaseI (Sigma, #D4527) or (Roche Cat#04716728001). Cell mixture was centrifuged at 1200pm (300g) for 5 minutes at room temperature. Supernatant including fat phase was removed and cell pellet was subjected to re-suspension in 1ml of Red Blood Cell Lysis Buffer (Sigma, #R7757) for 1 minute. Afterwards, the solution was re-suspended in 2% Tet-free FCS containing PBS (PF2) and cells were counted in preparation for the cell sorting and *in vitro* morphogenesis assay, and injections into the cleared fat pad.

2.3 Cell labeling, Flow Cytometry, and Sorting

After the red blood cell lysis buffer treatment, the number of cells obtained from each sample was counted and subjected to stain with a cocktail of antibodies. The antibody solutions were prepared according to the number of cells that are obtained from the previous procedure. The buffer used for the antibody staining and washing steps was PF2 (PBS plus 2% FBS). Firstly, the cell pellet were subjected to blocking with rat anti-CD16/CD32 for 5 minutes, then washed and pelleted. Lineage of lymphocytes, erythrocytes and endothelial cells were stained with a cocktail of purified CD45, Ter119, CD31 (all 1:100 from eBioscience) antibodies for 5 minutes in order to separate from mammary epithelial cells. After washing in PF2, cells were stained with PE-Cy5 F(ab')2 fragment of goat anti rat IgG (H+L) (Invitrogen) that detect all anti-rat CD45, Ter119, CD31 antibodies. After washing in PF2, cells were stained with anti-CD24-eFluor450 (1:100 eBioscience), anti CD29-APC/Cy7 (1:100 Biolegend) and anti-CD49f-PE (1:800 Millipore) for 5 minutes at room temperature. CD29, CD24, CD49F is a combination of published stem cell markers from Shackleton *et al.*[118] and Stingl *et al.*[128]. We used two additional cell surface antibodies CD1d (1:50 PE eBioscience) for the enrichment of more refined stem cell population, and CD61 (1:800) for the enrichment of the progenitors in both luminal and myo-epithelial cell populations[132]. Following washing with PF2, cells were re-

suspended to the 10,000cells/µl suitable for cell sorting. 7-Aminoactinomycin D, (7-AAD) (Sigma, #A9400) 1µg/ml as added 5min prior to FACS analysis for exclusion of dead cells. Samples were analyzed and sorted using a flow cytometer FACS Aria (BD Biosciences) at a concentration of 300cells/sounds and cell populations of interest are collected into 1.5ml eppendorf DNA low binding tubes containing 500µl of growth media. Data were analyzed using the FlowJo (version 8.8.7) software.

2.4 *in vitro* 3D organotypic culture

FACS sorted cells were collected in 500µl of culture media and pelleted immediately after sorting. The cells were re-suspended in growth factor enriched MEGM medium (Lonza/Amaxa MEBM Basal Medium (Lonza/Amaxa Cat#CC-3151 with added MEGM Bulletkit Cat#CC3150, plus additional B27, bFGF, hEGF). The cells were plated in 10,000-20,000cells/100ul of cell density in 50% of matrigel, 50% of media. The12 well plates or special glass-slide bottomed chambers, Lab-Tek II Chambered Coverglass (M-Medical S.r.l #FA9155379) were used for culture and future microscopic analysis. The maximum cell density was 200cells/µl. Co-culture gels of 50,000 unsorted cells were placed also on the side of the each experimental matrigel for the first 5-7 days of culture in order to increase the survival rate of cells after FACS.

The droplets of matrigel containing cells were solidified at 37°C in a CO_2 incubator for 45-60 minutes prior to addition of 1.5ml supplemented serum free MEGM (Mammary Epithelial Cell Medium BulletKit, #CC- 3150, containing one 500mL bottle of Mammary Epithelial Cell Basal Medium and supplements: 2 mL of bovine pituitary extract,, 0.5 mL of hEGF, 0.5 mL of hydrocortisone,, 0.5 mL of GA-1000, 0.5 mL Insulin, [Cambrex]). First 3-7 days of culture, the enriched media supplemented with B27, bFGF, hEGF were used. The culture media was changed every other day.

2.4.1 *in vitro* tumorigenesis assays

We induced oncogene overexpression under the control of the doxycycline inducible Tet-On system with tissue specific promoter, MMTV-rtTA. The *in vitro* transgene expression was achieved through the addition of doxycycline (Sigma) at a final concentration of 1ug/ml to the culture media. After FACS and plating of cells, we let the cells settle in the *in vitro* culture environment for at least 24 hours before inducing with doxycycline. The cellular response to oncogene overexpression varied between the cell types and also the type of oncogenes.

2.4.2 Passaging the organoid in 3D culture

Matrigel containing cells was pre-digest in Collagenase 3 and Liberase TM for 2 hours. The gel with media was transferred to 15ml falcon and washed with PBS. After centrifugation, the cells were re-suspend in 200µl of 0.25% trypsin EDTA (per 100µl gel) for 10 minutes in 37°C. Trypsin was inactivated with 10% Tet-Free Serum containing media and with 5-15mg/ml DNAseI. After centrifugation, cells were re-suspend in media followed by mixing with 50% of matrigel and allowing it to polymerize it in the incubator.

2.5. Vaginal Smears

Following the modified protocol of McLean *et al*. (2012)[155], vaginal Smears were performed using 20µl-40µl of PBS, and allowed to dry on the slide at room temperature. Once dried, smears were fixed in 10% formalin for 1 minute, 10% PBS wash for 1 minute, subsequently stained with Crystal Violet solution (Sigma) for 1 minute, followed by a one-minute wash in running tap water. Slides were dried and visualized on a Leica LMD 7000 mounted with Leica CD310 digital camera using LASV3.7 (Leica) software.

2.6 *in vivo* cleared fat pad mammary gland reconstitution assay

The sorted cells obtained from FACS were pelleted and re-suspended in an appropriate cellular concentration for the further transplantation assay. 3 weeks old female $Rag1^{-/-}$ mice were anesthetized and the mammary anlage of the inguinal mammary glands were dissected. The rest of the fat pad that has no mammary gland became the site of transplantation. The volume of injection was 10µl using a hamilton syringe (Hamilton Company, (83700) PB600-1 dispenser, (81101) 250 µL, Model 1725 LT SYR, NDL). The cells were prepared shortly before the injection re-suspended in 50% of matrigel 50% of media.

2.7 *in vivo* tumorigenesis assay

Tumorigenesis assays were preformed via injection of serial dilution of FACS enriched cells into the inguinal cleared mammary fat pad of $Rag1^{-/-}$ mice. Upon transplantation, the recipient mice were fed doxycycline containing food 625 mg/kg, or 0.012mg/ml in drinking water for 3 days low dose induction were specified in order to activate the transgene. Once the developed tumors in inguinal mammary fat pads were detected and reached a maximum 1 cm in diameter, the tumor bearing and the

same experimental cohort animals were sacrificed, and the tumors were fixed in formalin prior to embedding.

2.8 Transplantation *in vivo* after culture

Similarly to the previously described method, 2.4.3. Passaging the organoid/colonies in 3D culture, single cell suspension was obtained and after counting, the cells were mixed with matrigel (50%) shortly before injection and transplanted into 3 weeks old cleared fat pad of $Rag1^{-/-}$ female mice.

2.9 Immunofluorescence of *in-vitro* 3D culture whole mount

The cells cultured in Lab-Tek II Chambered Coverglass (M-Medical S.r.l #FA9155379) were subjected to immunostaining and further microscopic analysis. Prior to the fixation with 4% PFA, the matrigel containing the cells and colonies were partially digested by collagenase and liberase (3µl each in 1.5µl culture media) for less than 30 minutes. The enzymes were washed out with PBS for 5min 37°C once. The matrigel containing the cells was washed once with PBS and fixed in 4% Paraformaldehyde (PFA). Subsequently, after three times of washing with 1X IF buffer 5 minutes each, the matrigel was incubated with primary block (1X IF buffer + 10% goat serum Jackson Immunoresearch #005-000-121) for 1 hour. 10X IF buffer was made of 38g NaCl, 9.38g Na2HPO4, 2.07g NaH2PO4, 2.5g NaN3, 5g BSA, 10ml Triton-X100, 2.05ml Tween-20 in 500ml total volume in DDW (pH 7.4, filter sterilized). To make 1X IF buffer, 10X IF buffer was diluted in DDW. We made another incubation with 1ml of secondary block (primary block + 1:100 dilution of AffiniPure F(ab')2 Fragment Goat Anti-Mouse IgG, F(ab')2 Fragment Specific 1mg/ml stock, Jacksoon Immunoresearch #115-006-006) for 30min in order to reduce the unspecific binding and background staining as much as possible. Overnight incubation with primary antibody diluted in secondary block was routinely used in order to fully penetrate the antibody throughout the matrigel at 4C with gentle agitation. The primary antibody was washed with IX IF buffer three times for 30 minutes each, and subsequently incubated with secondary antibody diluted in primary block for 1 hour. DAPI (4',6-diamidino-2-phenylindole) was also added at this step. The secondary antibody solution was washed away with IX IF buffer three times for 30 minutes each and rinsed with PBS three times for 10 minutes each. At the end, the matrigel containing cells was mounted with VECTASHIELD® Mounting Medium with DAPI (Vinci-Biochem # VC-H-1200-L010).

Alternatively, the matrigel of 3D mammary cultures can be embedded in paraffin. Cultures were washed with PBS prior to fixation. 4% PFA was added and incubated for 10 minutes at room temperature. After three PBS washes, the gels were subjected to dehydration steps as following; 50% EtOH wash, 70% EtOH, 80% EtOH, 96%, EtOH 100%. It was stained with Eosin, and again 100% EtOH, twice Xylene, and twice paraffin at 56˚C followed by paraffin embedment. Paraffin blocks were cut at 3µm. The section can be subjected to immunohistochemistry or immunofluorescence as following.

2.10 Histochemistry, Immunohistochemistry, and Immunofluorescence

2.10.1 H&E

Tissues were fixed in 10% formalin overnight and washed in PBS prior to dehydration in Leica ASP300S, followed by embedding in paraffin in accordance with standard procedures. Tissue paraffin blocks were sliced at 5µm. Tissue sections were placed onto SuperFrost® Plus glass slides. De-paraffinization of section was carried out as follows: 15 minutes incubation in Xylene twice, 3 minutes 100% EtOH three times, 3 minutes 96% EtOH, 3 minutes 70% EtOH, 2 minutes DDW for 5 minutes. Sections were stained for haematoxyline for 1 minute and washed in running water. Subsequently, the slides were counterstained with hematoxylin then tissue sections were dehydrated and mounted. Dehydration steps are in 70% EtOH for 3 minutes, 96% EtOH for 3 minutes, 100% EtOH twice 3 minutes, then xylene twice 3 minutes each. We used DPX mounting media DPX mounting media (VWR #360294H). Images were obtained on a Leica LMD 7000 mounted with Leica CD310 digital camera using LASV3.7 (Leica) software.

2.10.2 Immunohistochemistry on paraffin sections

After the hydration step of paraffin embedded sections described above, they were incubated in antigen retrieval solution for 40 minutes at 100˚C (Vector Lab Antigen Unmasking solution Vector #H-3300). Following cooling (50˚C) endogenous peroxidases were inactivated via incubation in 3% H2O2 for 10 minutes. Following thorough PBS washing for twice 5 minutes each section was blocked with 10% Goat Serum (Vectashield ABC Kit) diluted in PBS and incubated for 30 minutes. Primary antibody was diluted in blocking solution at specified concentrations in table below (Table 5.). Sections were incubated with primary antibody at 4˚C overnight or 1-2 hours at room temperature. Sections were washed with PBS for 5 minutes twice. We

used Vector Lab Biotinylated Secondary antibody kit for secondary antibody incubation (30 minutes at room temperature). Subsequently, sections were washed for 5 minutes twice with PBS. Following the instructions of Vector Lab ABC reagent we added this to the sections and incubated for 30 minutes at room temperature. Following washing as described before the peroxidase reaction was carried out using Vector Lab DAB kit (#SK-4100). Sections were incubated for less than 1 minute and as soon as the stain was visible, the slides were washed in DDW. The slides were counterstained with hematoxylin then tissue sections were dehydrated and mounted. Images were obtained on a Leica LMD 7000 mounted with Leica CD310 digital camera using LASV3.7 (Leica) software.

Table 5. List of antibodies used in immuno-stainings.

Name of Antibodies (concentration used)	**Species of the antibody host**	**Supplier**
Actin (Smooth Muscle) Monoclonal Mouse Anti-Human (1:500)	Mouse	Dako
TROMA-I cytokeratin Endo-A conc. (1:250)	Rat Polyclonal	DSHB
Keratin 14 Polyclonal Antibody, Purified (1:1000)	Rabbit	Covance
Monoclonal Anti-SMA 1A4 (1:500)	Mouse	Sigma
Neu Antibody (C-18) (1:200)	Rabbit Polyclonal	Santa-Cruz
c-Myc (D84C12) Rabbit mAb (1:400)	Rabbit	Cell Signaling
Slug (C19G7) Rabbit mAb #9585 (1:100)	Rabbit	Cell Signaling Technology
Cleaved-Caspase-3 (1:200)	Rabbit	Cell Signaling Technology
Green Fluorescent Protein (GFP) Antibody (1:200)	Chicken	Aves Lab
GFP antibody (ab290) (1:200)	Rabbit	Abcam

2.10.3 Immunofluorescence on paraffin sections

Deparaffinization, hydration and antigen retrieval was carried out as for IHC sections. Sections were blocked using 10% goat serum in 1X IF buffer (as previously described) for 1 hour at room temperature. Sections were incubated with primary antibodies for 1-2 hours at room temperature. Sections were washed three times in 1X IF buffer for 10 minutes each. Alexa-Fluor secondary antibodies (invitrogen) were incubated at 1:800 (Goat anti-rat-647, anti-rabbit 568, Goat anti-mouse 488) with DAPI 1:1000 for 1 hour at room temperature. Subsequently, they were washed three times with X1 IF buffer (10 minutes each), and three times in PBS (5 minutes each), mounted with MOWIOL® 4-88 Reagent (Merk #475904) and stored at 4˚C.

2.11 Carnoy's fixation and whole mount carmine alum staining

Mammary glands were dissected out and spread on the glass slide (SuperFrost® Plus glass slides). After up to 30 minutes of drying the tissue in the room temperature, it was fixed in Carnoy's fixative (60% of absolute EtOH, 30% of CHCl3, 10% of glacial acetic acid) for 1 hour or overnight at room temperature. Before staining, the MG was washed in 70 % EtOH and change gradually to distilled water (70%, 50%, 0%) for 15 minutes each. MG was then stained in Carmine Alum Staining solution (Bio Optica #05-B07009) for 1 hour or overnight depending on the thickness of the tissue. The tissues were subsequently washed in acidic alcohol (70% EtOH, 1% HCl) to reduce excess of staining for overnight incubation. The tissue was dehydrated in 96 % EtOH 30 minutes, 100 % EtOH 30 minutes and then in xylene until the fat pad was be transparent. In the end, the tissue was mounted with DPX mounting media (VWR #360294H). Imaging was carried out using Stereomicroscope (Leica DC500).

2.12 Real-Time Quantitative PCR

Total RNA was extracted from cells using Trizol (Invitrogen). cDNA preparation, including DNase digestion, was performed using QuantiTect Reverse Transcription kit (Qiagen). Reverse transcription reaction was completed for each sample; 200ng of mRNA was converted to cDNA in each RT reaction. Amplifications were run using technical triplicates and biological triplicates in a LightCycler 480 (Roche). Values were adjusted using Actin-B as a reference. The table below shows the list of primers used in the experiments.

Table 6. List of primers for qRT-PCR

Gene	Forward	Reverse
HER2	TGTACCTTGGGACCAGCTCT	GGAGCAGGGCCTGATGTGGGTT
MYC	AGATGGTGACCGAGCTGCTGG	AAGCCGCTCCACATACAGTC
L37a	TCTGTGGCAAGACCAAGATG	GACAGCAGGGCTTCTACTGG
Actin B	GCTTCTTTGCAGCTCCTTCGT	ACCAGCCGCAGCGATATCG
GAPDH	CCCATTCTCGGCCTTGACTGT	GTGGAGATTGTTGCCATCAACGA

Primers were purchased from Sigma Aldrich. qPCR reaction was carried using the following concentrations; Primers at 5μM, Syber Green master mix, 14ng of cDNA were used for each reaction. Actin B primers were used to normalize results.

2.13 EdU incorporation

Click-iT® EdU Alexa Fluor® 488 Imaging Kit (Life technologies #C10337) was used following manufactures instruction. EdU incorporation was performed 18 hours before fixation of the matrigel. After the staining procedure, images were acquired by confocal microscope, resonant scanner or spinning disk microscope and analyzed with velocity software (Perkin Elmer).

3. Results

3.1 Expression of Myc- versus Neu- oncogenes in adult mammary gland leads to development of primary mammary tumors with contrasting tumor phenotypes.

As previously shown in Moody *et al.* and D'cruz *et al.*, we achieved overexpression of oncogenes with the new generation of tractable transgenic mouse models, Tet-On-Neu/MMTV-rtTA (T-O-NEU) and Tet-On-c-Myc/MMTV-rtTA (T-O-MYC), in the adult mammary gland to better mimic the situation in the clinic[148,156]. The oncogenes were induced and overexpressed after virgin female T-O-NEU and T-O-MYC mice reached their adulthood at 8 weeks of age. This is an alternative to the constitutive models, which overexpress the transgene from the embryonic stage throughout life. The tetracycline inducible system enables us to control the overexpression of specific genes at any given time point, in this case Neu (murine form of HER2) and c-Myc[52]. We detected palpable primary mammary tumors in T-O-NEU and T-O-MYC animals with an average of 3 and 6 months of induction, respectively (Figure 9. Data obtained by Martina Mantovan).

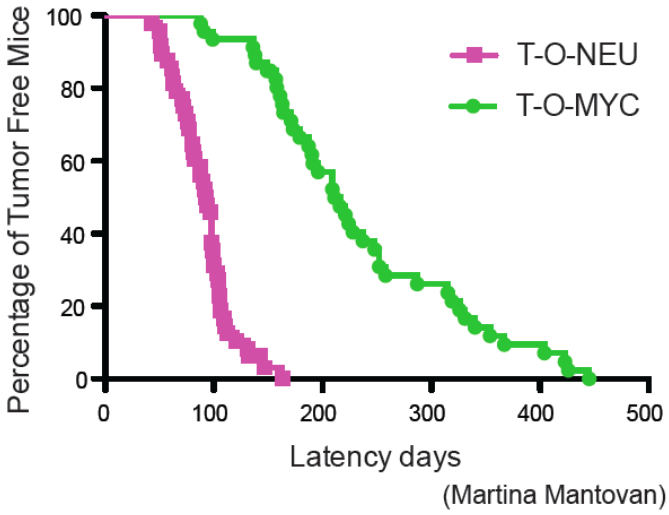


Figure 9. The administration of doxycycline to the adult female T-O-MYC and T-O-NEU mice leads to the development of primary mammary tumors with contrasting tumor phenotypes. Primary tumor latency of T-O-NEU (n=49) and T-O-MYC (n=45) mice induced with doxycycline after 8 weeks of age. Neu overexpression leads to the development of palpable tumors after a median latency of 3 months, while in Myc expressing mice the median latency was 6 months. (Data produced by Martina Mantovan)

The mammary gland is composed of bi-layered ducts, the outer layer consists of myoepithelial/basal cells and the inner layer is built with luminal cells (Figure 10). The myoepithelial lineage can be visualized by the staining of smooth muscle action

(SMA) or other cytokeratin markers such as cytokeratin 14 (K14). In contrast, the luminal lineage expresses high level of cytokeratin 8 and 18 (K8/18). These two distinct mammary cellular lineages are also used to characterize the mammary tumor arising from T-O-NEU and T-O-MYC mice.

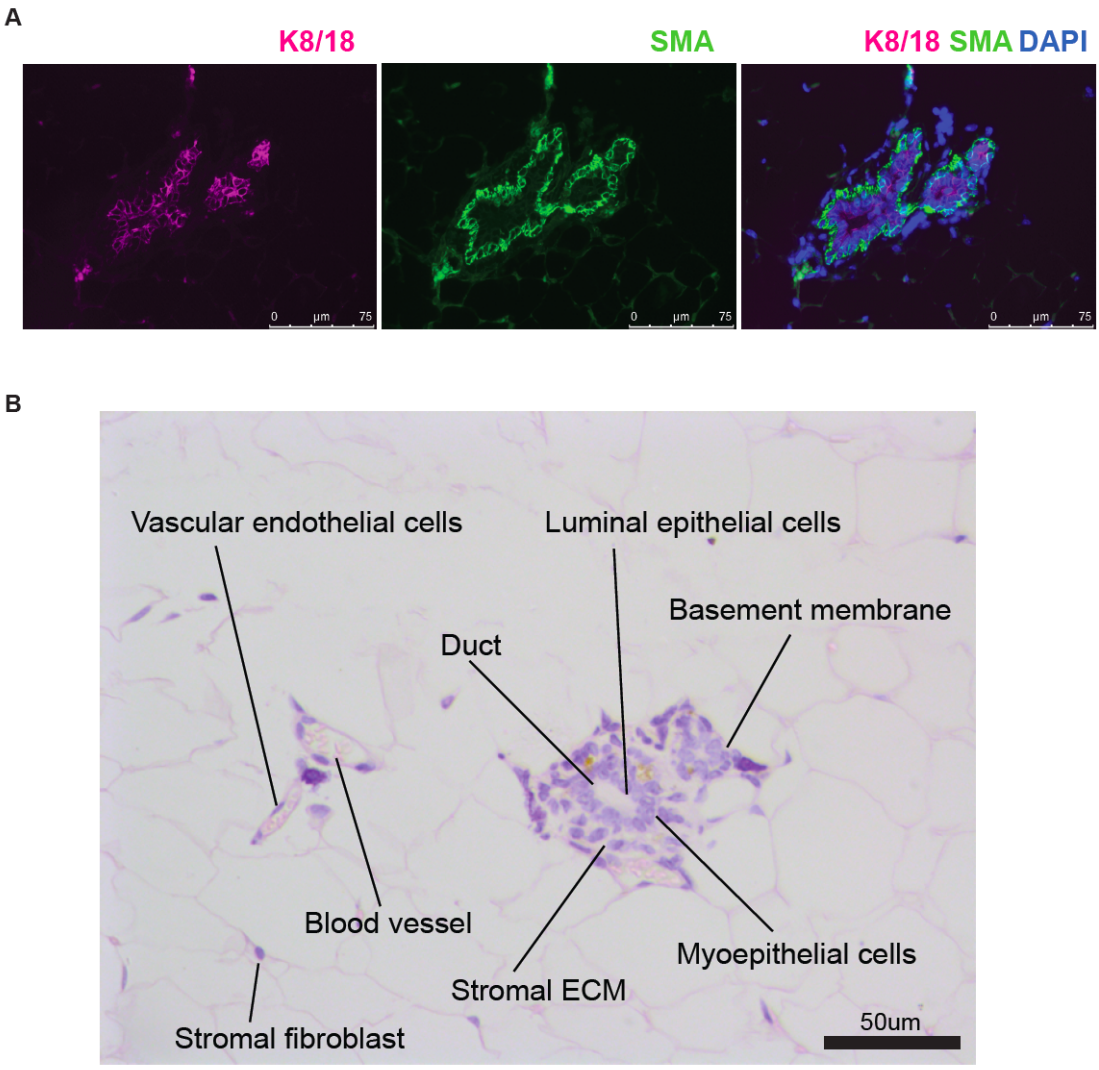


Figure 10. Mammary gland is composed of bi-layer duct of Luminal (K8/18) and Myoepithelial (SMA) lineage surrounded by the stroma. A) Immuno-fluorescence staining of an adult wild type mouse mammary gland (age 8 weeks) stained with SMA (myoepithelial/basal marker) and K8/18 (luminal/epithelial marker). This clearly demonstrates that the mammary gland is a bi-layered duct. B) H&E staining of adult mouse mammary glands showing the composition of the mammary epithelia and stroma.

As previously described[157], Neu overexpression leads to the development of mammary tumors consisting of mainly luminal-like epithelial cells and display the classic cobblestone-like morphology (Figure 11. A-C). All tumors were of the solid phenotype, and the majority of tumor cells expressed luminal marker K8/18. In contrast, c-Myc driven tumors are reported as basal phenotype[37], which give rise to

heterogeneous tumor subtypes[93,158]. The immuno-staining of the lineage specific markers K8/18 and SMA confirmed that c-Myc overexpression leads to basal type mammary tumors, which are composed of both myoepithelial cells and luminal cells (Figure 11. D-F).

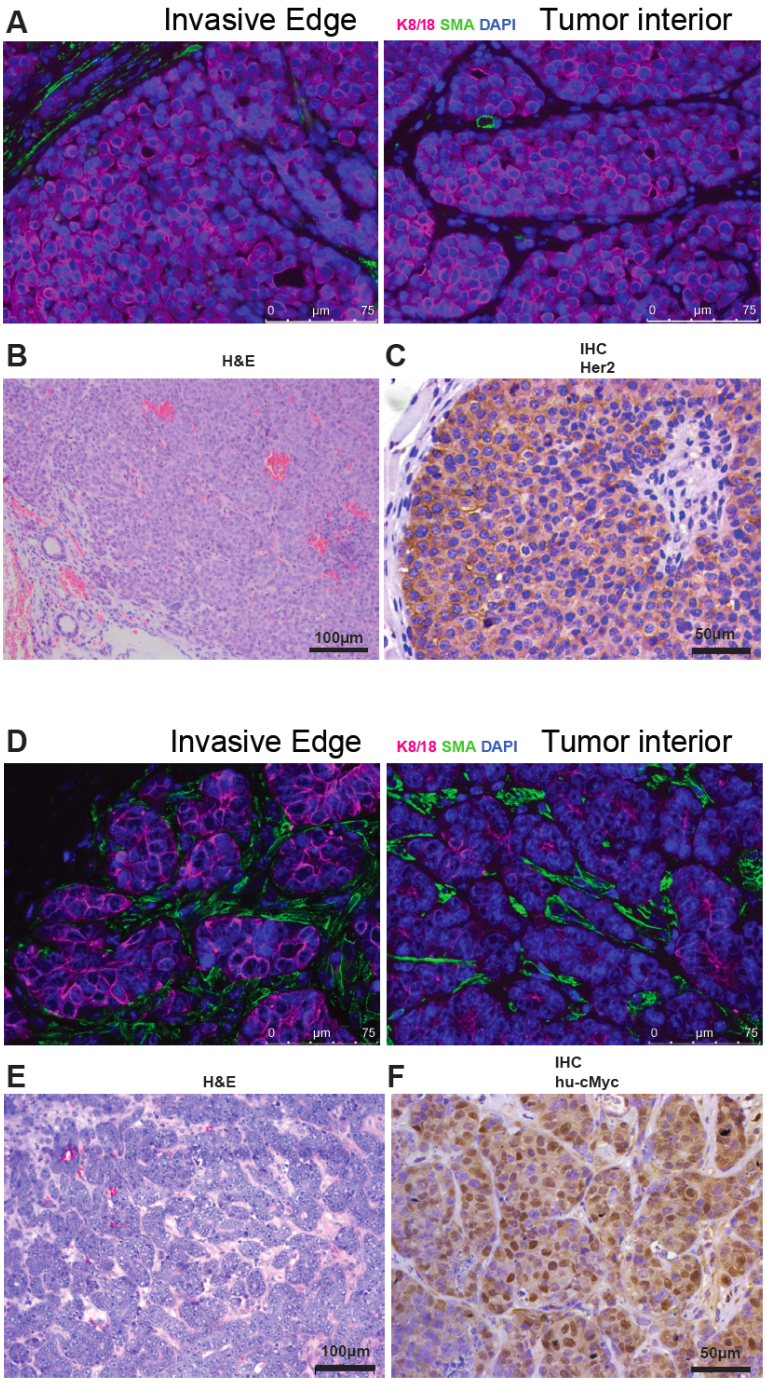


Figure 11. Two oncogenes, Neu and Myc, give rise to tumors that are histologically distinct. A) Immuno-fluorescence staining of Neu induced primary tumors by K8/18 and SMA showing Neu-driven tumors are composed of mainly luminal cells. B) H&E representation of Neu tumor (20X). C) Immuno-histochemistry of HER2 staining (40X). D) Immuno-fluorescence staining of Myc induced primary tumors by K8/18 and SMA. This shows Myc drives tumors to be a mixed/basal phenotype. E) H&E representation of c-Myc tumor (20X). F) Immuno-histochemistry of human-c-Myc staining (40X).

3.2 *in vitro* cultures recapitulate normal development of mammary epithelial cells

In order to address the question, which cell types are the TICs upon overexpression of Neu or Myc, we first took an *in vitro* approach. Using a 3D matrix culture system, we initially aimed to validate the characteristics of the different sorted cell populations and to observe the cellular transition from healthy adult mammary epithelia to the early tumorigenic state (Scheme in Figure 8 in introduction). In order to separate different cell populations of interest within the mammary linage, we employed FACS (Fluorescence-Activating Cell Sorting) according to the expression levels of specific cell surface proteins[118,128].

To date, the mammary cellular hierarchy has not been fully characterized, however, a number of studies suggest the presence of defined cellular compartments. By staining mammary epithelial cells with a combination of antibodies detecting unique cell surface markers, we could achieve the enrichment of the adult bi-potent stem cells, the luminal progenitors, the basal lineage as well as the stromal population. In more detail, naïve mammary glands from virgin, adult female bi-transgenic (T-O-MYC or T-O-NEU) mice, that have never been fed doxycycline containing food and therefore never experienced transgene expression, were dissociated and digested into a suspension of single cells. The mammary cell suspension was subjected to a staining procedure by the combination of the antibodies that are directly conjugated with the specific fluoro-chromes.

In order to separate hematopoietic lineages including erythrocytes, lymphocytes, endothelial cells, we also included the linage exclusion step, by staining with CD31, CD45, Ter119 antibodies. A combination of cell surface proteins including CD29, CD24, and CD49f was used to distinguish mammary luminal and basal lineage as well as enrichment of luminal progenitors and the stem cells[118,128]. Both CD29 (b1 integrin) and CD49f (a6 integrin) are integrin subunits, which serve as heterodimeric extracellular matrix receptors. We included both antibodies in our FACS gating strategy in order to confine our gating strategies (Figure 12).

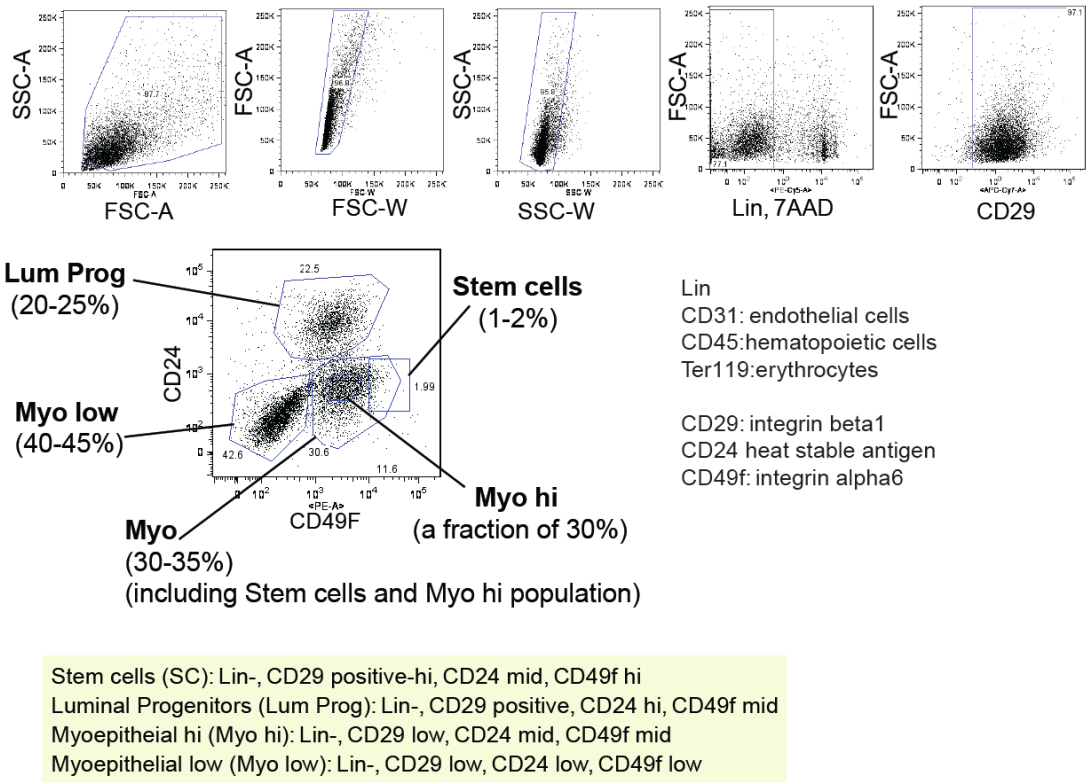
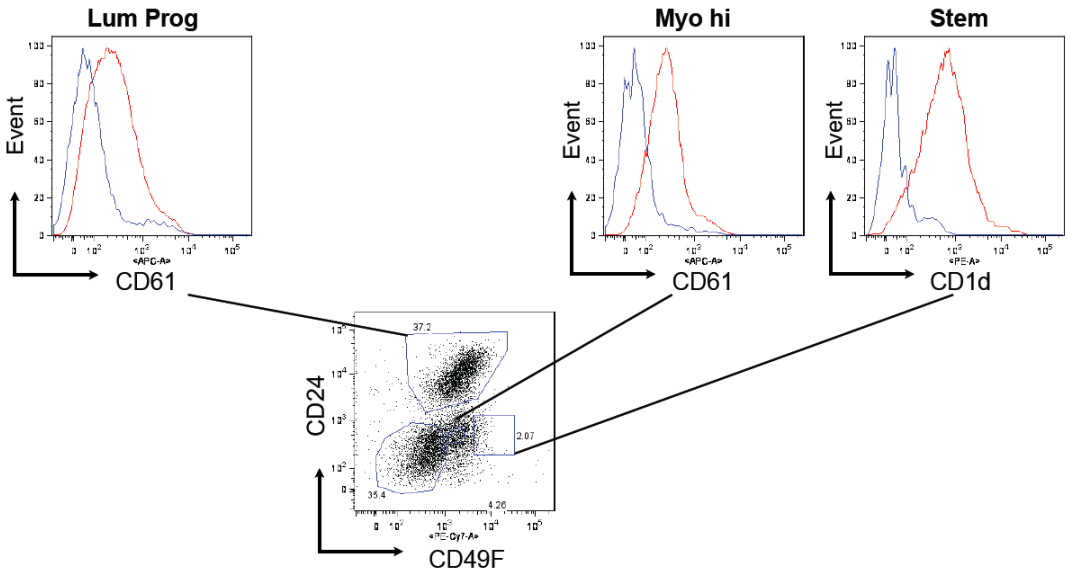


Figure 12. FACS gating strategies to isolate different mammary cellular lineages. Briefly, after the elimination of doublets, Lin (lineage) positive and 7-aminoactinomycin D (7-AAD) positive cells are excluded. Lin- population were gated into CD29 positive and separated into CD24 versus CD49f for the separation of stem cells, luminal progenitors, Myo hi and Myo low populations.

We define our "stem cell" population to be CD29$^{positive-hi}$, CD24^{mid}, CD49f^{hi} and gate for the 1-2% of highest CD49f expressing cells among the entire Myoepithelial population (CD29^{positive}, CD24^{mid}, CD49f^{positive} around 30% of total) (Figure 12). The myoepithelial population is known as a mixed myoepithelial population including the bi-potent stem cells, myoepithelial stem cells, myoepithelial progenitor cells and fully differentiated myoepithelial cells. We defined a middle population within the myoepithelial population as the "Myo hi" (CD29$^{low-positive}$, CD24^{mid}, CD49f^{mid}). By contrast, the luminal cellular lineage can be enriched by CD24 high expression. We defined the population CD29^{positive}, CD24^{hi}, CD49f^{mid} population (20-25% of total) as "Luminal Progenitors" (Lum Prog). However, this population also includes all the luminal lineage committed uni-potent stem cells, luminal progenitor cells, and differentiated luminal alveoli and ductal cells. Lastly, we have categorized the CD29$^{low-positive}$, CD24^{low}, CD49^{low} population as "Myo low". This population contains differentiated myoepithelial cells or sometimes also defined as fibroblast like stromal cells[129].

Since these well-defined markers can enrich for partially mixed populations, we attempted to further purify the specific cell populations with additional cell surface markers. Previously, comparative gene expression profiling of FACS enriched stem cells versus other mammary epithelial populations showed no evident differences[128]. However, a recent RNA sequencing study by dos Santos *et al.* suggested CD1d to be the marker of more refined stem cell population[132]. In this study, CD61, Integrin β3, are also reported to enrich for progenitor cells or uni-potent stem cells of both luminal and myoepithelial lineage.

We attempted to combine these two cell surface proteins in order to improve our FACS gating strategies. We compared upper/lower 30% of CD1d expressing cells in the conventional stem cell gate, and upper/lower 30% of CD61 expressing cells in the luminal progenitors and myoepithelial population (Figure 13).



(CD61 hi Prog; CD61: integrin beta-3, and CD1d:MHC-protein that are involved in the presentation of lipid antigen to T cells.

Figure 13. FACS gating strategies to further refine the stem and progenitor populations using CD1d and CD61 antibodies. The FMO control (in blue) and the stained sample (in red).

Upon isolation of these FACS isolated cell populations, we attempted to keep them in culture and to expand the different populations while maintaining their own cellular characteristics. We have modified the previously described three-dimensional (3D) organotypic culture condition[141] and successfully established *in vitro* primary culture condition where FACS sorted cells are able to proliferate longer than 3 weeks in basement membrane extract (BME) matrix. Some stem cells colonies started to form

bi-layered acini after 1 week in culture, but many stayed as single cells or slowly proliferated and formed small cell aggregates. Without this optimization, we observed no colonies growing out within the culture, even from the luminal progenitor population (data not shown). This tells us that the FACS procedure has a detrimental effect on cellular proliferation. However, we found four key factors that overcome this effect from FACS and allow the cells to proliferate in culture. Firstly, we grow the sorted cells in presence of a co-culture of 50,000 unsorted mammary cells for the initial first week of culture. Secondly, we found that addition of supplements to the culture medium (EGF, FGF, and B27 for initial 3-4 days of culture) often used in the mammosphere assay increased the number of colony formations. Thirdly, the FACS sorted cells were usually plated in a concentration of 150-200cells/µl in 100µl volume. When culture contained a lower concentration, the majority of the cells did not proliferate (data not shown). Lastly, we tested the optimum concentration of BME and found that 50% (7,5mg/ml BME) would give the maximum number of colonies in both stem and luminal progenitor cells (Figure 14. A-B). A BME concentration of lowers than 7,5mg/ml is more prone to cause gel destabilization during long-term culture. Before optimization, using the protocol from Jechlinger *et al.,* (2009)[141], we recorded less than 1% of colony formation in all populations including stem (0.053%), luminal progenitor (0.46%) and Myo hi (0.026%). After the optimization of the culture condition in the 3D matrix, the colony forming efficiency of these sorted populations in culture was increased to 4.17% in stem (<50µm in diameter), 16.57% in luminal progenitor cells (hollow polarized acini >50µm), and 4.3% in Myo hi population (<50µm in diameter) after 7 days in culture (Figure 14. C).

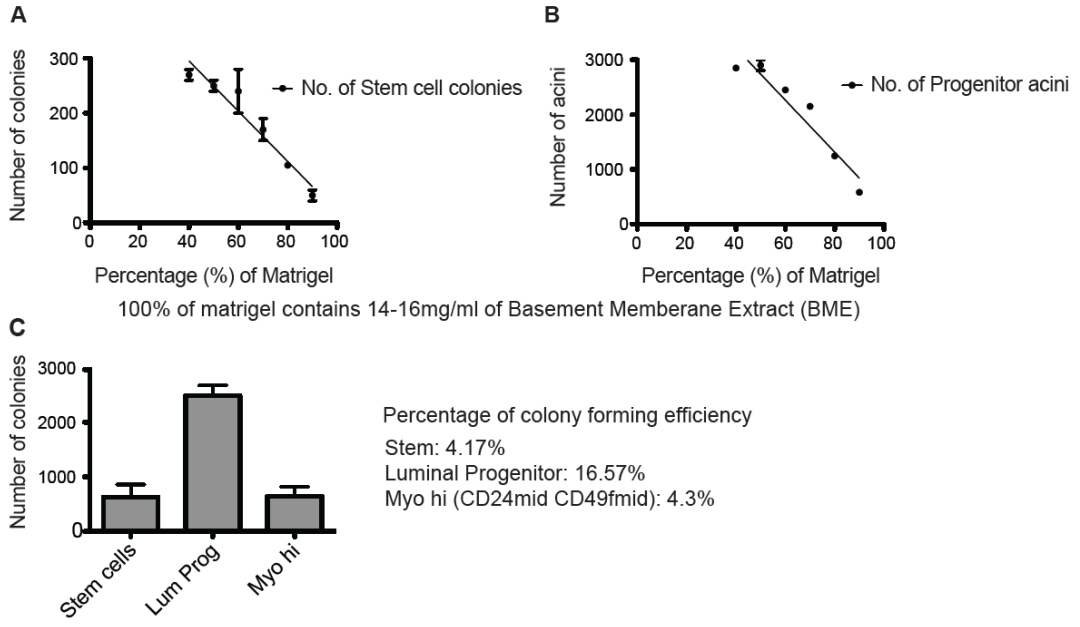


Figure 14. Optimized culture condition allowed the stem cells, luminal progenitors, and Myo hi population to expand and form colonies. A) Colony forming efficiency of stem cells can depend on the culture conditions including the consistency of BME. B) Colony forming efficiency of luminal progenitor cells can depend on the culture conditions including the consistency of BME. C) Colony forming efficiency of different sorted populations in optimized 3D culture (7d in culture of 15,000 cells/100µl). Each condition was repeated at least 3 times.

The Myo hi population gave rise to a similar number of stem-like filled colonies, however, this population was a mix of stem-like colonies and stromal-like cells that resemble the Myo low population. Interestingly, we observed a significant increase (nearly 20 times) in the number of acini formation in $CD61^{hi}$ upper 30% when compared with the $CD61^{low}$ lower 30% (Figure 15). CD61 expression in other populations including stem, Myo hi and Myo low population showed no significant differences (data not shown).

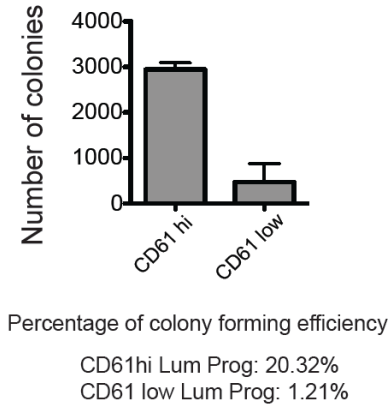


Figure 15. Significant difference was observed in number of colonies (luminal acini) arising from $CD61^{hi}$ (30%) and $CD61^{low}$ (30%) expression luminal progenitor cells. CD61 enriched the highly proliferating luminal progenitor cells (7d in culture of 15,000 cells/100µl)

However, CD1d expression did not show a significant difference from the bulk of stem cells or even less colony formation in culture. This hints that $CD1d^{hi}$ stem cells are more quiescent stem cells (data not shown).

The single cells obtained after FACS were plated in matrigel and developed colonies or hollow acini within the 3D matrix. These structures in the matrix seemed to maintain cellular characteristic features. After a week in culture, stem cells ($CD29^{hi}$, $CD24^{mid}$, $CD49f^{hi}$) grew into small filled colonies (approximately 50um in diameter) that co-express both K8/18 and SMA at the single cell level (Figure 16. A Left upper and Figure 16. B). Co-expression of both the luminal and myoepithelial lineage indicates that these stem cells are still bi-potent maintaining their differentiation

capacity. In contrast, the luminal progenitor cells (CD29^{mid}, CD24^{hi}, CD49f^{mid}) formed larger hollow polarized acini (approximately 100µm in diameter) that predominantly express the luminal markers, K8/18 (Figure 16.A Right upper). The Myo hi population was a mixture of stem cell like colonies as well as stromal or fibroblast like cells (Figure 16. A Left lower). Using an additional antibody, CD61, in the myoepithelial population, we could not observe any significant enrichment of myoepithelial progenitor cells (data not shown). Lastly, the Myo low population contained mainly the stromal cells (Figure 16.A Right lower).

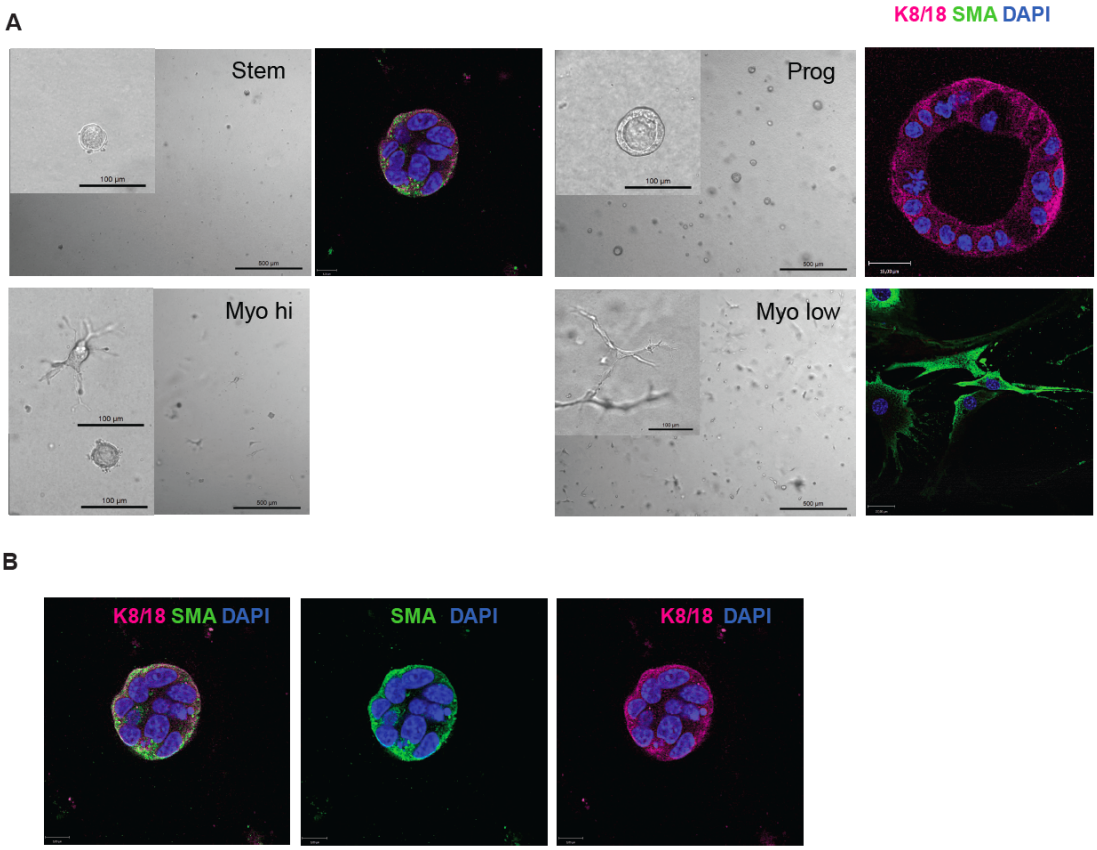


Figure 16. FACS enriched population give rise to colonies or acini and express lineage specific proteins. A) Bright field images in culture and co-immuno-staining of K8/18 and SMA demonstrate that the FACS enriched cells were able to proliferate in 3D culture maintaining their cellular lineages (7d in culture of stem cells, upper left, luminal progenitors in upper right, Myo hi population in lower left, and Myo low population in lower right). B) Co-immuno-staining of K8/18 and SMA demonstrates that the stem cell colonies co-express both luminal and myoepithelial lineage specific markers.

Interestingly, all of the luminal progenitor acini developed hollow acini that express the luminal marker, K8/18, however, some acini also co-expressed SMA moderately (Figure 17. A). This finding infers that luminal progenitor cells are composed of at least two different subsets within their hierarchy. Using the putative progenitor

marker, CD61, we try to further purify the bulk of luminal progenitor population into CD61^{hi} and CD61^{low} expressing populations. From the bright field microscopic analysis of acini in culture, in addition to an increase in the number of acini, an increased thickness of the rim of acini was noticed among the CD61^{hi} Luminal progenitors (Figure 15, Figure 17. B). After 1 week in culture of WT luminal cells, CD61^{hi} showed greater number of colonies with mild SMA positivity compare to CD61^{low} counterpart (Figure 17. C Left). When the culture time was extended to almost double (13 days) there was a decrease in the percentage colonies that express weak SMA and these barely showed any difference between CD61^{hi} and CD61^{low} luminal populations (Figure 17. C Right).

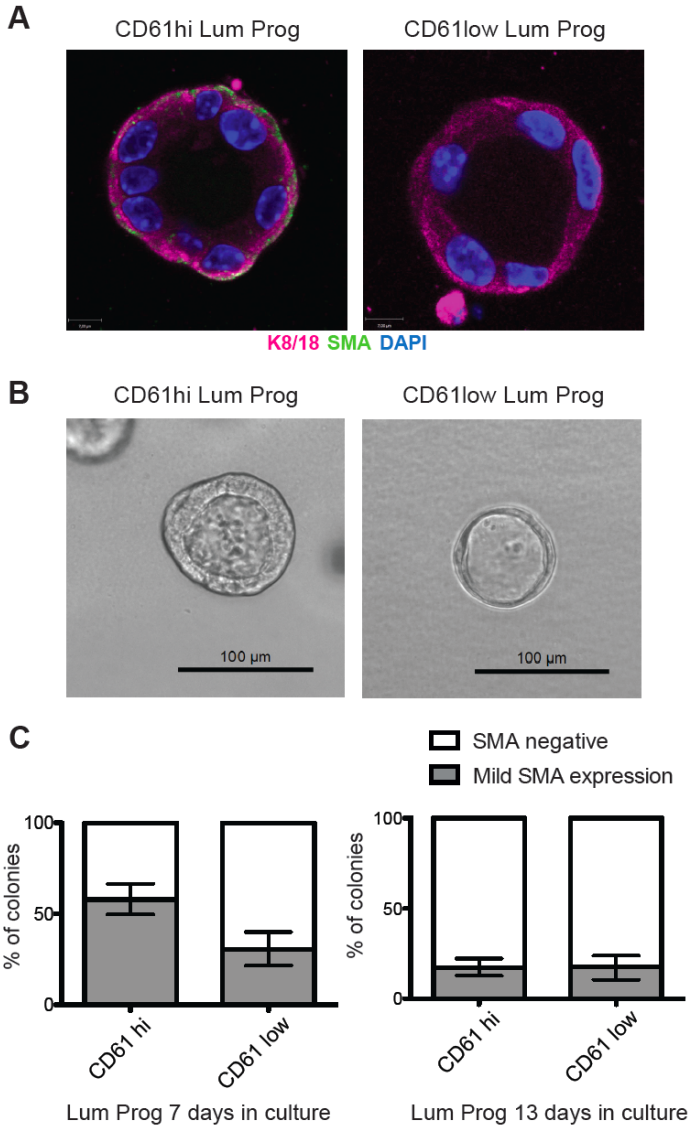


Figure 17. Mind SMA expression among CD61^{hi} Luminal progenitors became insignificant over time. A) Representative images of SMA positive and SMA negative luminal progenitor acini. B) Representative bright field images of colonies arising from CD61^{hi} and CD61^{low} luminal progenitors. C) Percentage of luminal acini that co-express K8/18 and SMA after 7 days and 13 days in culture. Decrease of SMA expression in CD61^{hi} Luminal progenitor acini were noted after over 10 days of culture.

These *in vitro* observations (Figure 15 and 17) support the previous report indicating CD61 as the marker of progenitor or transient amplifying cells within the luminal lineage. Moreover, it encourages that most of the luminal progenitor acini grow out from the un-refined Lum Prog ($CD29^{mid}$, $CD24^{hi}$, $CD49f^{mid}$) are the CD61 positive true luminal progenitors. This implies the presence of positive selection towards luminal progenitor cells by our *in vitro* culture condition.

In addition to the cellular lineage specific markers such as SMA and K8/18, we attempted to verify epithelial polarization status of FACS enriched stem cells in culture. Since the stem cell colonies after 1 week in culture appeared to be filled and we could not distinguish if the small colonies we observed were just a crumb of cells or if they had some kind of cellular organization; hence the epithelial polarization sttus was investigated. Previously, Jechlinger *et al.* reported that the acini growing out in 3D culture are composed of polarized cells with a correct orientation[141]. Zona Occludens-1 (ZO1), a protein that directly binds to occludins and marks tight junctions, is a known tight junction marker. The immuno-staining with ZO1 identified that stem cells gave rise to polarized colonies in culture with the apical surface juxtaposed to the lumen (Figure 18. A-B). As a positive control for correct epithelial polarization, we stained the luminal progenitor acini.

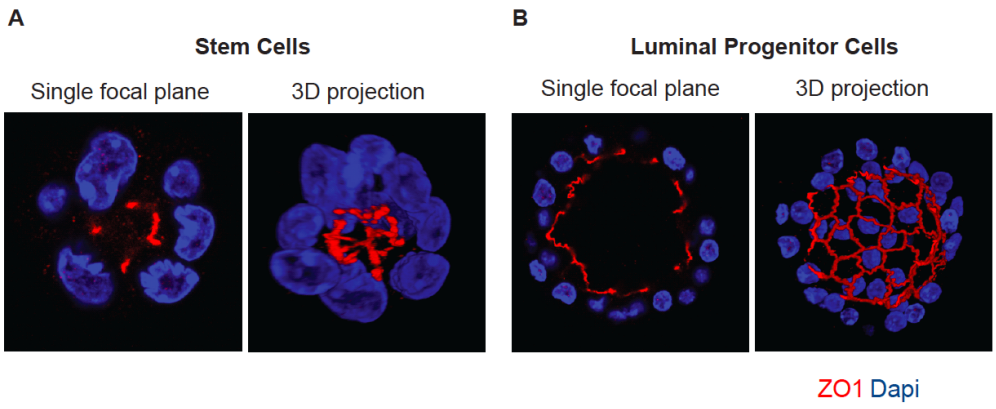


Figure 18. Both stem cells and luminal progenitor cells maintain their correct epithelial polarity. Immuno-staining of ZO1 (apical marker) shows the epithelial polarization take place in both A) stem cell colonies and B) luminal progenitor acini in culture.

3.3 FACS enriched stem cells can reconstitute the mammary gland upon transplantation

In addition to the *in vitro* analysis, we assessed if the sorted stem cells in the culture after 1 week would still maintain their cellular characteristics and would be able to repopulate a functional mammary gland upon transplantation. The cultured stem cell colonies and luminal progenitor acini in culture were isolated by digestion of the matrigel with Collagenase 3 and TM Liberase and dissociated into single-cell suspension upon treatment with 0.25% trypsin EDTA for 5 minutes. Serial dilutions of these cultured cells were subjected to the cleared mammary fat pad *in vivo* transplantation assay, a gold standard assay to assess the stem-ness by reconstitution of a functional mammary gland[117] (Figure 6 introduction).

Before examining the mammary gland reconstitution potential of cultured stem cells, we first performed the *in vivo* transplantation assay using FACS enriched stem cells without the culture process. Single cell suspension of FACS isolated stem cells as well as other cell populations were individually transplanted in serial cellular dilution. The site of transplantation was the fat pad of the inguinal mammary glands of 21days old female mice that were surgically removed of their endogenous mammary gland. The fat pads of recipient mice were analyzed after 4-12 weeks in order to assess the mammary gland reconstitution (Figure 19). 4 weeks after the transplantation, a mammary gland with immature duct that filled approximately 30% of entire fat pad was observed. 12 weeks after transplantation, the reconstituted mammary gland filled the entire fat pad and the mature ducts resembled th endogenous mammary gland with no transplantation. In higher concentrations, the FACS enriched luminal progenitor cells were also able to reconstitute the mammary gland, however, it filled less than 30% of entire fat pad.

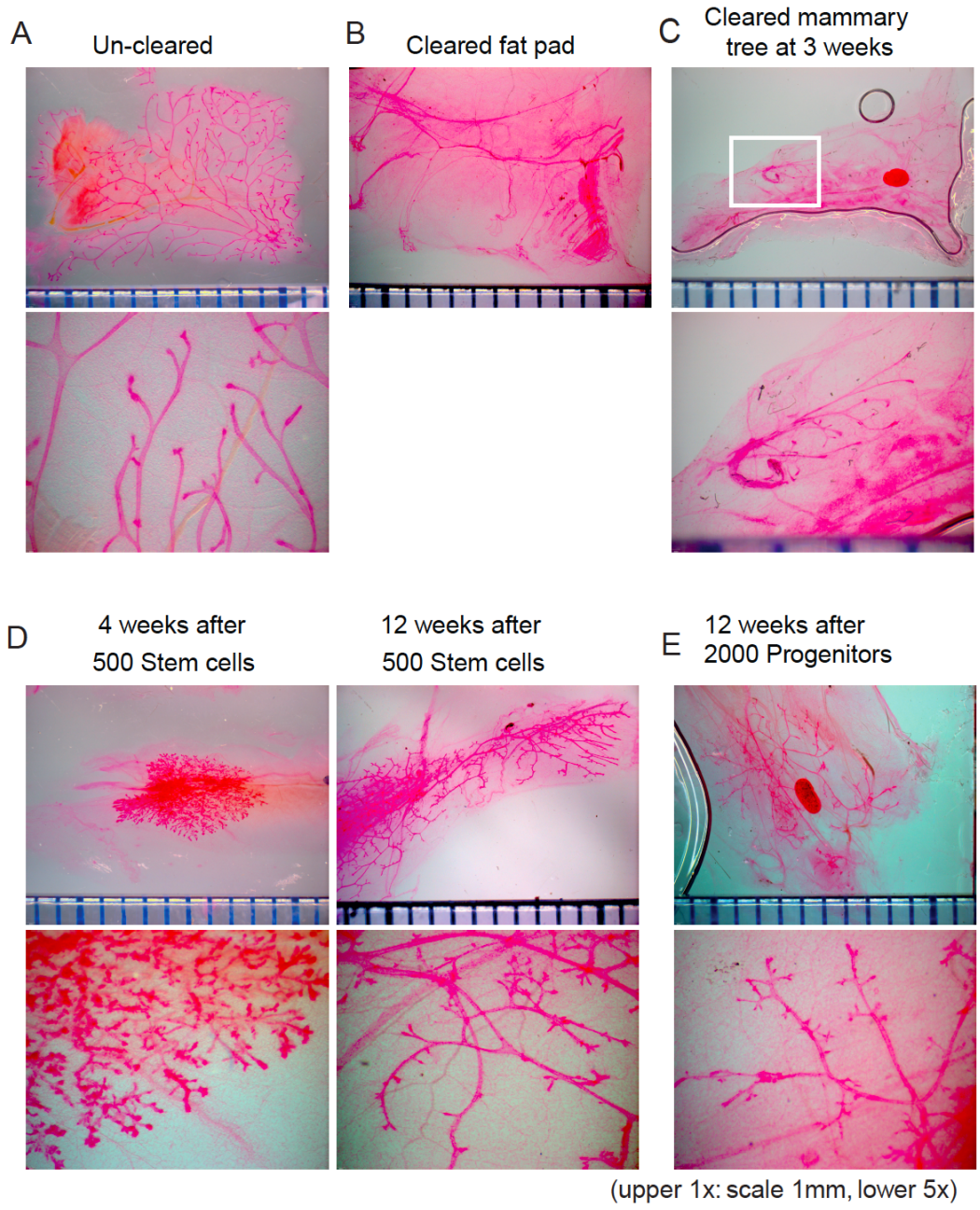


Figure 19. Whole mount carmine alum staining of reconstituted and the control mammary glands. Controls: A) Representative images of a cleared fat pad at adult stage. B) Representative image of the cleared fat pad at 3 weeks of age. C) Representative images of the cleared mammary anlage surgically removed when the recipient mice are at 3 weeks of age. D) Representative images of the mammary fat pad 4 weeks and 12 weeks after injection of 500 stem cells. E) Representative image of the mammary fat pad injected with 2000 luminal progenitors. (Upper panel 1X and lower panel 5X images).

The estimated frequency of mammary repopulating units (MRU) was 1/196 in stem and 1/1614 in luminal progenitors calculated by Extreme Limiting Dilution Analysis, ELDA[138] (Figure 20). However, while stem cells were able to reconstitute almost the entire fat pad (>80%) with a mammary tree, luminal progenitor cells gave rise to reconstituted mammary trees that filled less than 30% of the fat pad (Figure 19 E).

Compared to previously published MRU estimations (1/64 from Shackleton *et al.,* 2006[118]; 1/60 (in FVB background) and 1/90 (in C57B1/6 background) from Stingl *et al.,* 2006[128]), our result showed 50-33% of their published MRU efficiency. This could be due to internal variability and setting of the experiment and the fact that Rag1$^{-/-}$ mice retain some level of immune function including natural killer (NK) cell activity. Also, previous studies describe this assay utilizing syngeneic mice with a pure genetic background. Our transgenic mice are bred in a mixed background (between C57Bl/6N, FVB/NHanTM Hsd, and 129S2/SvHsd), therefore, we decided to use female Rag1 knock out mice in C57Bl/6N background as the transplantation recipient.

	Number of cells injected per fat pad	Number of outgrowths	Fat pad filled	Repopulating frequency
Stem cells	2000	7/7	● x7	1/196
	1000	2/2	● x2	
	500	5/7	○ x3 ● x2	
	100	3/4	○ x1 ◕ x2 ● x1	
	50	2/4	○ x2 ◕ x1 ● x1	
	10	0/1	○ x1	
	5	1/1	◕ x1	
Luminal Progenitors	2000	4/10	○ x6 ◔ x3 ◕ x1	1/1614
	1000	2/2	◔ x2	
	500	3/7	○ x4 ◔ x3	
	100	1/5	○ x4 ◔ x1	
	50	2/3	○ x1 ◔ x2	
	10	0/1	○ x1	
	5	0/1	○ x1	
Myo low	2000	0/4	○ x4	1/9748
	1000	-		
	500	1/4	○ x3 ◔ x1	
	100	-		
	50	-		
	10	-		
	5	-		

○ No reconstitution of mammary gland

◔ Reconstituted mammary gland filled 1-30% of fat pad

◕ Reconstituted mammary gland filled 31-80% of fat pad

● Reconstituted mammary gland filled 81-100% of fat pad

Figure 20. Table displays the result of transplantation WT FACS enriched populations. Full-fillness of the fat pad and their repopulating efficiency (MRU: Mammary Repopulating Unit) are presented from Stem, Luminal Pogenitors and Myo low populations.

The reconstituted mammary glands were further analyzed by immuno-staining with SMA and K8/18. Upon injection of 500 stem cells, the reconstituted mammary glands were clearly composed of an outer layer of myoepithelial cells that are positive for the SMA staining, and the inner layer of luminal cells that are positive for the K8/18

(Figure 21. A Left). This result indicates that the reconstituted mammary gland has the same functional structure as that of the normal adult mammary gland. Interestingly, luminal progenitor cells were also able to reconstitute bi-layered mammary gland though higher cellular dilutions were required (Figure 21. A Right). We attempted to induce pregnancy and lactation in these reconstituted mammary glands. These data, although preliminary, both stem cells and luminal progenitors derived reconstituted mammary gland responded to pregnancy, documented indication of proliferation. Again, the presence of both luminal and myoepithelial cells were confirmed in the mammary gland subjected to pregnancy (Figure 21 B). From this finding, de-differentiation of luminal progenitors upon transplantation was suspected. Luminal progenitor cells were able to reconstitute functional mammary trees although the efficiency and the area of fullness were less than that of stem cells.

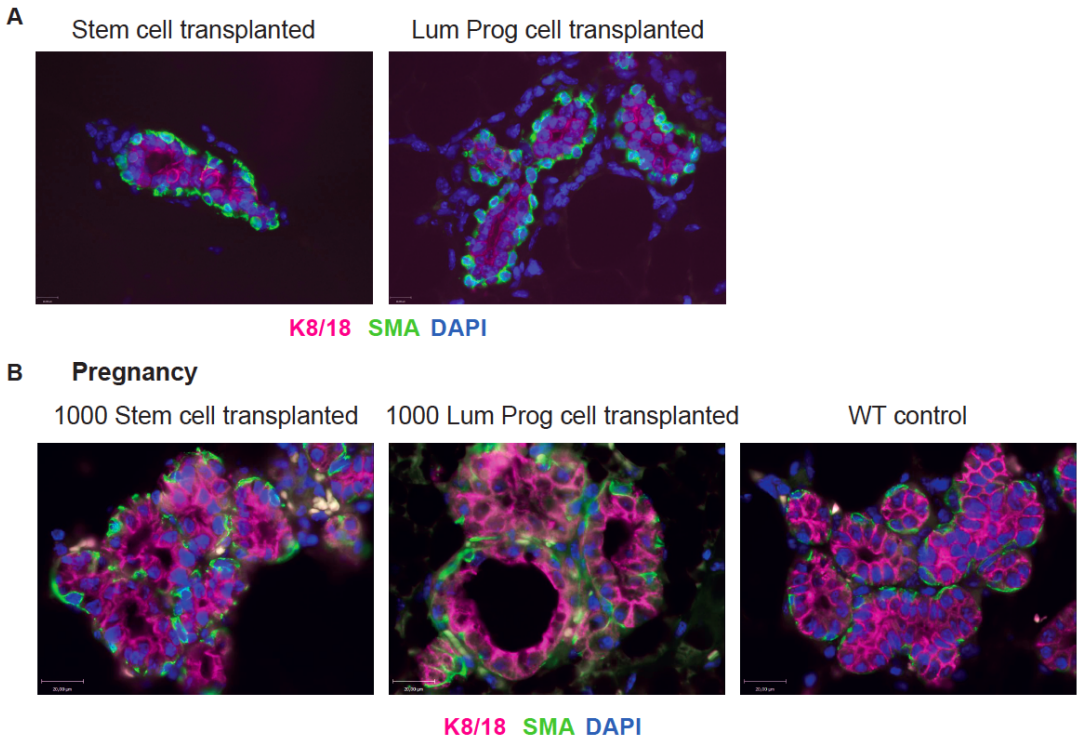


Figure 21. Reconstituted mammary gland consisted of bi-layer of luminal (K8/18 positive) and myoepithelial (SMA positive) cells. A) Left: image of stem cell derived reconstituted mammary gland, Right: image of luminal progenitor cell derived reconstituted mammary gland. B) Recipient mice were mated and induced pregnancy. Both stem cells (Left) and luminal progenitor (Middle) derived mammary glands showed expansion of a bi-layered mammary tree resembling control. (Right) normal un-transplanted mammary gland induced pregnancy.

Since we had lower frequency of MRU, we also tried to refine the purity of stem and luminal progenitor population further employing two additional cell surface antibodies, CD1d, an early marker for stem cells, and CD61, a marker of lineage committed progenitors[132]. Against our expectation, we could not confirm the enrichment of

refined stem cells as previously reported[132]. The *in vivo* result was not conclusive that CD1d enriches stem cells, or CD61 enriches progenitor cells within the myoepithelial lineage (data not shown).

3.4 FACS enriched and then cultured stem cells can reconstitute the mammary gland upon transplantation

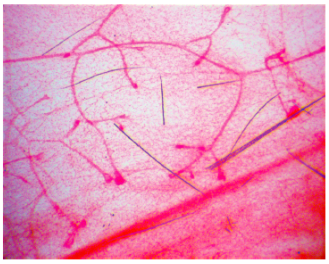
Using the above explained *in vivo* transplantation approach, we attempted to assess the stem cells that were kept in culture if they could still maintain the stem-ness (i.e. being able to reconstitute the mammary gland by self-renewal and differentiation). The cultured colonies were dissociated and single cell suspension was obtained. Although in a lower efficiency, transplanted cells from stem cell colonies in culture for 1 week were able to reconstitute the mammary glands (Figure 22). The estimated MRU was 1/742 in cultured stem cells (Figure 22. A). We also examined the same in cultured luminal progenitors and its estimated MRU was 1/5005 in cultured luminal progenitors. The reconstituted mammary gland from the transplanted cultured stem cell injection filled the cleared mammary fat pad greater (>80%) than the cultured luminal progenitors (<30%) (Figure 22. A). This result was similar to the transplantation of FACS enriched stem and luminal progenitors without culture. The analysis of whole mount staining of reconstituted mammary gland demonstrated that the luminal progenitors reconstituted mainly the ducts that lack the terminal end buds (Figure 22. C). In order to assess functionality of the reconstituted mammary gland, we induced pregnancy and examined the composition of mammary tree. Interestingly both reconstituted mammary gland from stem and luminal progenitor propagated in culture could expand into bi-layered ducts composed of luminal and myoepithelial cells (Figure 22. D and E). Although it is preliminary, we observed that the number and frequency of myoepithelial cells per duct were higher in the cultured stem cell derived reconstituted mammary gland.

In short, we confirmed that FACS enriched stem cells were able to proliferate in culture maintaining their distinct cellular characteristics. In addition, we observed that luminal progenitor cells, although in lower frequency, were also able to do so. Altogether, we have successfully set up an *in vitro* culture system, which allow us to follow the transformation process of the sorted stem and progenitor cells upon overexpression of oncogenes (induced by doxycycline in the culture medium).

A

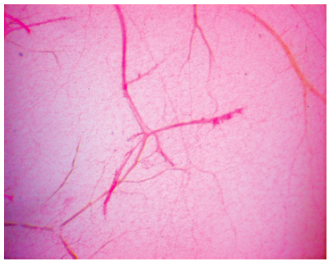
	Number of cells injected per fat pad	Number of outgrowths	Fat pad filled	Repopulating frequency
Cultured Stem cells	1000 500 100 50 10 5	1/4 2/4 2/2 1/5 1/3 0/2	○ x3 ◕ x1 ○ x2 ◕ x2 ◔ x1 ◕ x1 ○ x4 ◔ x1 ○ x2 ◔ x1 ○ x2	1/742
Cultured Luminal Progenitors	1000 500 100 50 10 5	1/7 1/6 0/5 0/5 0/2 0/2	○ x6 ◔ x1 ○ x5 ◔ x1 ○ x5 ○ x5 ○ x2 ○ x2	1/5005

B



C

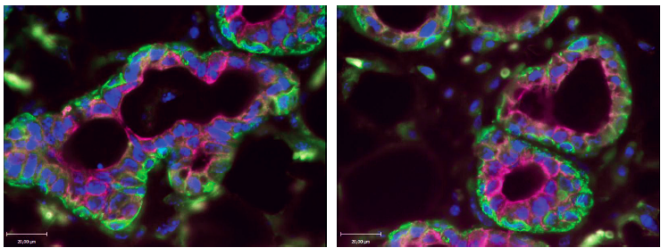
Cultured 1000 Lum Prog transplanted



Cultured 1000 Stem cells transplanted

D

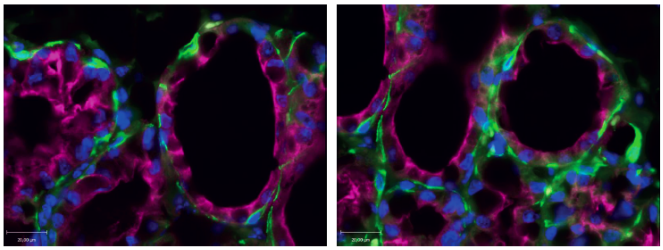
1000 Cultured Stem cells transplanted (Pregnancy)



K8/18 SMA DAPI

E

1000 Cultured Luminal progenitor cells transplanted (Pregnancy)



K8/18 SMA DAPI

Figure 22. Transplantation of FACS enriched stem and luminal progenitor cells after in culture for 1 week were able to give rise to functional mammary glands. A) Estimation of frequency of mammary repopulating units can be determined by using the ELDA. B-C) The whole-mount carmine alum staining of reconstituted mammary glands. FACS enriched and cultured stem cells are able to reconstitute the functional mammary gland with end buds but not in case of the luminal progenitor cells (0.8X and 5X images). D) The immunostaining of pregnancy induced reconstituted mammary glands from injection of stem cells. E) The immunostaining of pregnancy induced reconstituted mammary glands from injection of luminal progenitor cells.

3.5 *in vitro* cultures recapitulate early stage of Neu driven tumorigenesis *in vivo*

Using this organotypic *in vitro* 3D culture system, we monitored the proliferation of single sorted cells during early stages of tumorigenesis driven by the overexpression of two different oncogenes. Stem cells, luminal progenitors, Myo hi and Myo low population were obtained from the bi-transgenic mice T-O-NEU or T-O-MYC using previously described conventional FACS approach. After 24 hours of initial culture, in order to settle the sorted cells to its culture condition, doxycycline (1µg/ml) was added to the culture medium (onDOX-state) to express the transgenes Neu or Myc. In some populations, we found a significant difference in the proliferation and morphology of the colonies upon overexpression of the oncogenes. In order to establish that the differences between the onDOX and the Never induced state come from the activation of the transgenes, we first needed to confirm the transgene expression in each of the FACS enriched cell populations.

qRT-PCR analysis of the rat-Neu transgene showed that all FACS enriched populations overexpressed this oncogene 24 hours after induction (Figure 23. A). Transgene activation was consistently confirmed also in the later time points of the culture period (1, 2, and 3 weeks). After 3 weeks in culture, some stem cells and Myo hi population gave rise to morphologically aberrant larger colonies. We separately analyzed the expression of the Neu transgene in the larger colonies and the rest, but Neu transgene was overexpressed in all cases. The expression level was lower in Myo low (the stromal population) compared to the other three populations (Stem, Lum prog and Myo hi). Interestingly, the expression of the transgenes in the luminal progenitor acini decreased after 2 and 3 weeks in culture.

To confirm transgene expression at the protein level, we performed immuno-staining using an anti-HER2 antibody on the cultured colonies. Although this antibody recognizes both endogenous and exogenous expression of HER2, only some cells within the onDox state colonies appeared to express HER2 in the cellular membrane (Figure 23.B).

There was a clear Neu overexpression in cultured colonies shown at both the mRNA and protein level after 1 week on Doxycycline. However, we did not observe any significant morphological difference using the bright field microscopic analysis between the onDox and the Never state. After 2-3 weeks of culture, some filled

colonies arose from the stem and Myo hi population that showed moderate differences in morphology compared to their Never induced controls (Figure 23.C).

A fraction, 28.5% of colonies arising from the stem cells seem to respond to Neu oncogene overexpression by changing their morphology (Figure 23. C and Figure 24. A). The rest of stem cell colonies maintained their slower proliferation rate similar to their Never induced state. Similarly, 21.3% of Myo hi population also gave rise to colonies that resembled the ones from the stem cell population (Figure 23. C and Figure 24. B). Unexpectedly, no significant difference was observed in the luminal progenitor even after 3 weeks in culture onDox (Figure 24. C). The Myo low stromal population developed fibroblast like features in culture and there were neither the colony formation nor other noticeable morphological differences upon overexpression of Neu (Figure 24. D). In addition, further FACS enrichment of luminal progenitors into $CD61^{hi}$ and $CD61^{low}$ population in culture also exhibit no apparent morphological differences upon induction of Neu (data not shown). In addition, we did not observe any significant enrichment of Neu induced aberrant colonies within the $CD1d^{hi}$ stem cells or $CD61^{hi}$ Myo hi population (data not shown).

We next verified that the morphological changes in onDox-cultured colonies correlate with increases in proliferation. EdU incorporation confirmed that proliferation among the fraction of stem cell colonies was higher upon Neu overexpression (Figure 25. A-B). Along with the change in their morphology and proliferation rate, we also observed a gradual loss of myoepithelial lineage and expansion of luminal cells in those stem cell colonies (Figure 25. C). As mentioned above, the majority (over 90%) of stem cell colonies exhibited double positivity for SMA and K8/18 within a week in culture (Figure 16. B). After 2 weeks in culture, within the onDox stem cell cohort, the expansion of inner luminal compartment (K8/18 single positive cells) was observed. The outer layer of these colonies remained SMA positive varying in the degree of their expression. After 3 weeks of culture, SMA expression in the onDox stem cell colonies was very low or completely lost in some colonies, while expression of K8/18 luminal lineage was enhanced. These observations point towards the possibility that Neu overexpression in stem cells lead to tumor development by specifically contribute towards expanding the luminal-committed lineage.

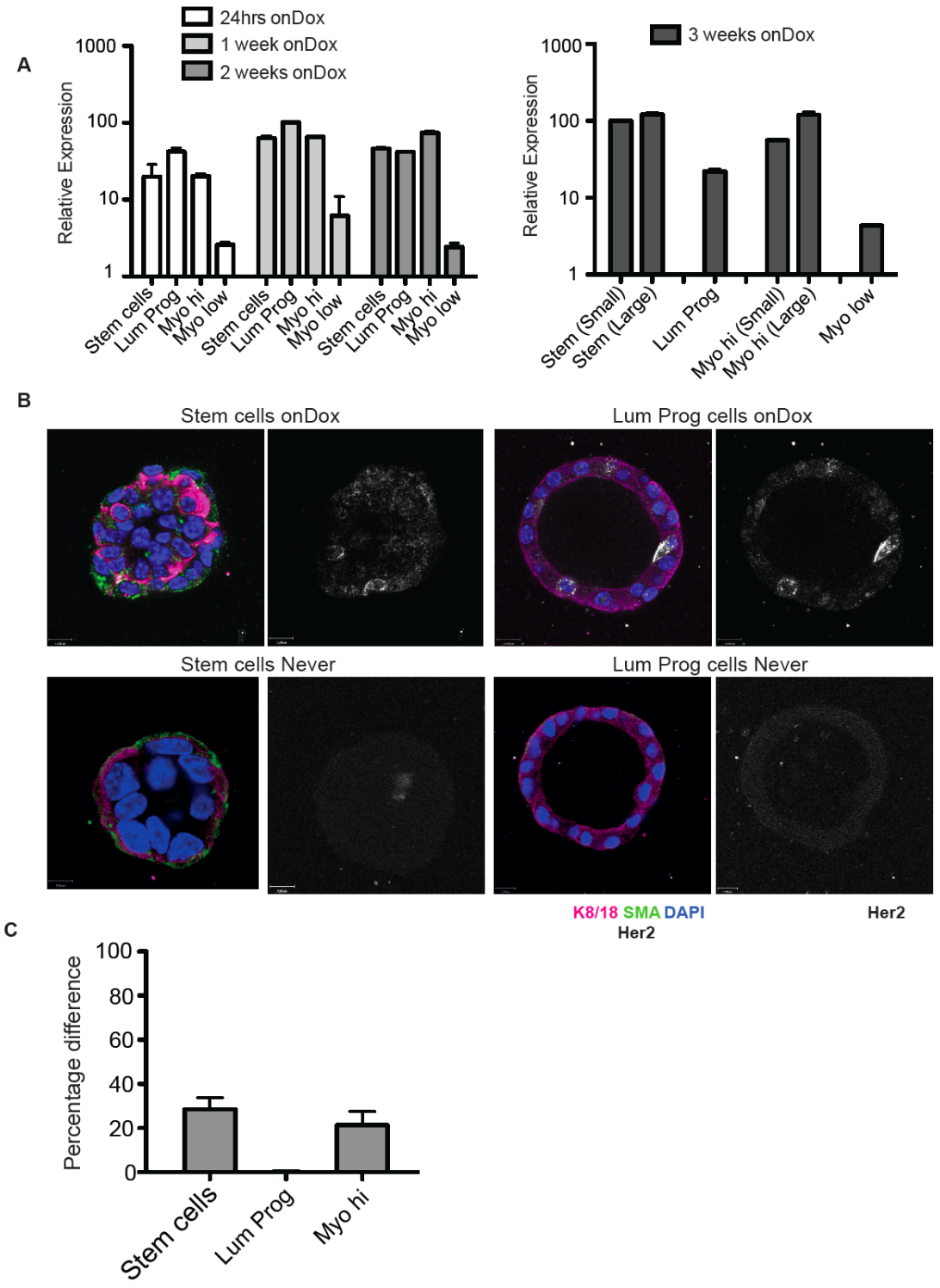


Figure 23. *in vitro* cultures recapitulate early stage of Neu driven tumorigenesis *in vivo*. A) Relative expression of rat-HER2 transgene in mRNA levels in the different FACS enriched cell populations detected by qRT-PCR. Y axis showed relative level of mRNA expression in logarithmic scale. Different fractions of FACS enriched cell population from T-O-NEU express the transgene rat-Neu in culture. B) Immuno-staining of the HER2 expression in cultured colonies (1 week in culture onDox). C) A fraction of induced stem cell colonies and Myo low colonies in culture exhibit significant morphological differences upon doxycycline induction compared to their never induced controls.

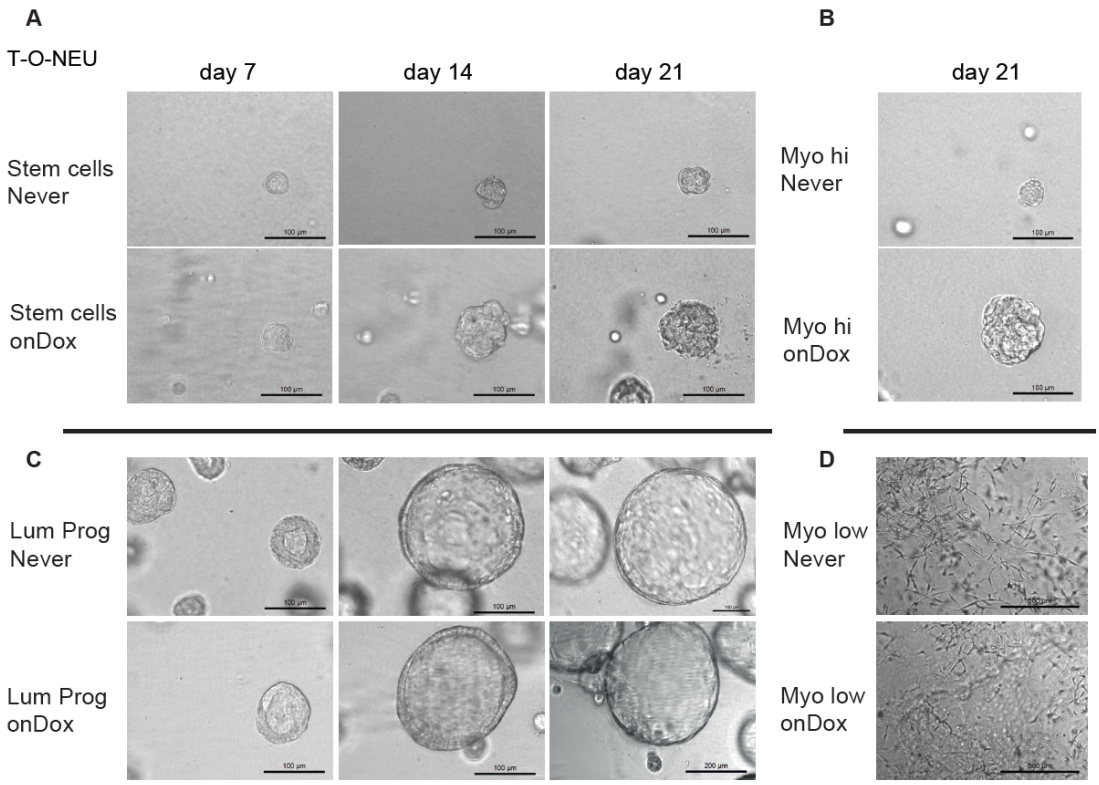


Figure 24. Bright field photo documentation analyses of the sorted cells from T-O-Neu in culture over 3 weeks of period. A) Representative images of stem cell colonies. A fraction of the induced stem cells (onDox) responds to transgene activation. B) Representative images of Myo hi colonies. Similar response was observed in Myo hi population. C) Representative images of luminal progenitor acini. No significant difference was seen in luminal progenitors. D) Representative images of Myo low population. No significant difference was seen in Myo low population.

Despite of the apparent transgene expression among luminal progenitors, we did not find any significant change in the morphology or proliferation rate upon doxycycline induction over 3 weeks of culture period. This lack of response among luminal progenitor cells *in vitro* might be explained by the contribution of the *in vivo* microenvironment in luminal progenitor cell driven mammary tumorigenesis, and would require further investigation. On the contrary, our *in vitro* culture hints that stem cells could well be TICs of Neu overexpressing breast tumors. Since Neu driven primary tumors are composed of mainly luminal cells (Figure 11. A), the preferential expansion of the luminal lineage observed in stem cell colonies in culture resembles the composition of the Neu primary tumors. Thus, our observation also adds stem cell population as an additional candidate tumor cell-of-origin in Neu driven tumors.

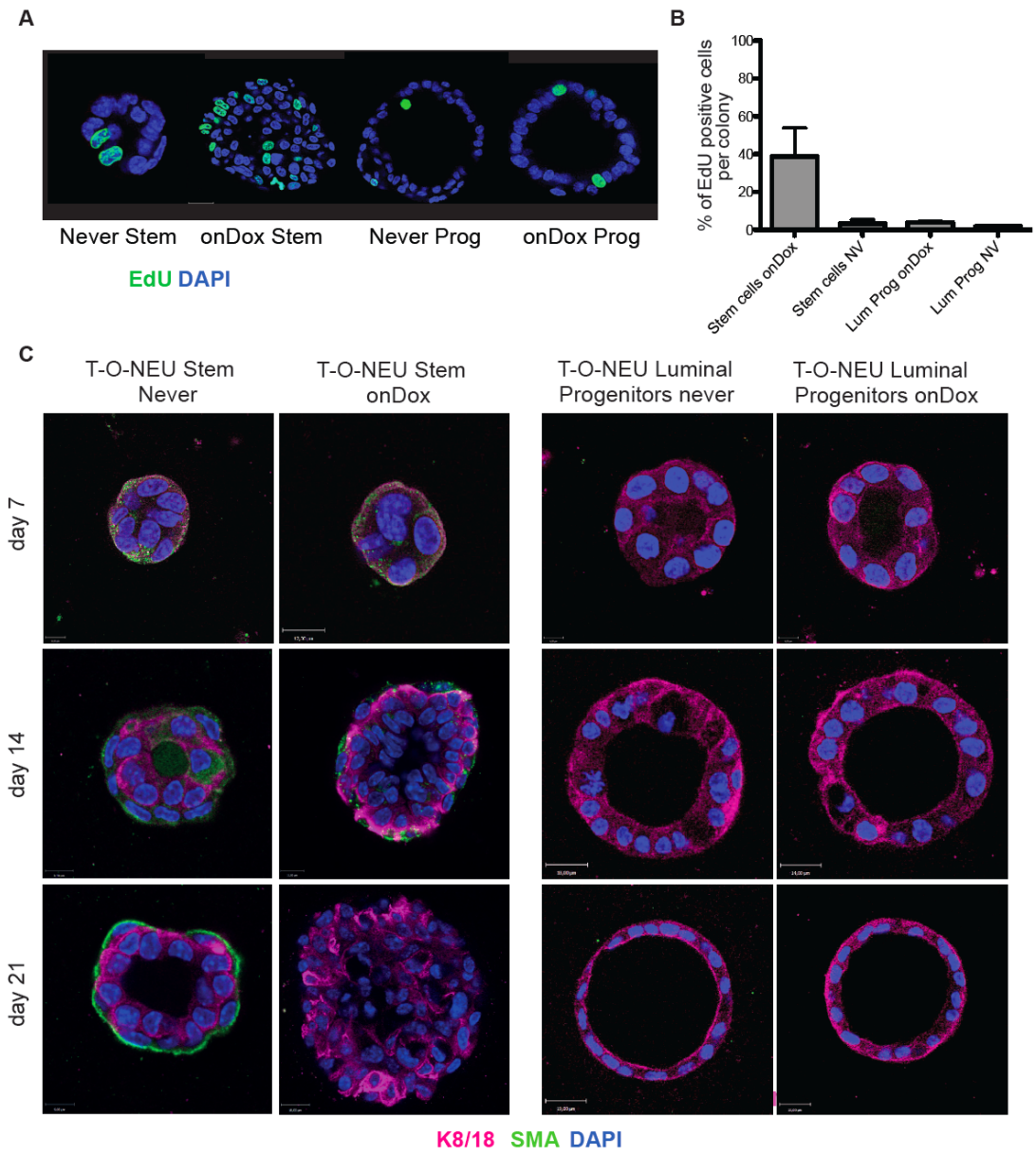


Figure 25. A fraction of stem cells responded to Neu overexpression in culture by increasing their proliferating and preferential differentiation towards luminal lineage. A) Representative pictures of Stem and Progenitor cells with and without doxycycline induction for 3 weeks, showing an increase number of positive EdU cells on the SC on dox. Green: EdU staining, blue: DAPI (EdU incorporation 18 hours no pause) B) Quantification of EdU incorporation in the spheres in culture (n>10 colonies). C) Immuno-fluorescence staining of cellular lineages (luminal and myoepithelial) of the cultured colonies. Stem cells colonies exhibit the expansion of the luminal lineage over a period of time.

3.6 in vitro cultures recapitulate early stage of Myc driven tumorigenesis *in vivo*

The Myc oncogene exhibited rapid morphological changes within the cultured stem cell derived colonies as well as luminal progenitor derived acini after 3-4 days onDox. In order to confirm if the c-Myc transgene is also expressed in all cell populations

obtained after FACS, we performed qRT-PCR analysis at the mRNA level of human c-Myc transgene after 24 hours and 1 week of culture (Figure 26. A). Transgene expression was detected in all cell populations including stem cells, luminal progenitors, Myo hi and Myo low population. We again separated morphologically aberrant larger colonies of the stem cell population and compared them to the rest. The Myc transgene was overexpressed in both cases. In contrast to Neu, there was a higher Myc transgene expression within the Myo low stromal population. Despite using the same promoter (MMTV) for the overexpression of the transgenes, we observed this difference in transgene expression. It implies that the c-Myc oncogene is preferentially expressed among the stromal cells.

In addition, we performed immuno-staining of human-c-Myc co-stained with SMA and K8/18 antibodies on stem and progenitor colonies in culture onDox for 7 days. This analysis clearly demonstrated the presence of transgene human-c-Myc expression at the protein level inside the nuclei (Figure 26. B). The Myc expressing cells in the stem cell colonies are normally the SMA positive myoepithelial cells and not the K8/18 positive luminal cells. In the progenitor cells Myc expression was present in both myoepithelial and K8/18 positive luminal cells or double-positive cells.

The c-Myc overexpression took a shorter time period (less than 1 week) in order to induce the apparent morphological cellular transformation of colonies in comparison to Neu (became significant only after 3 weeks onDox). The percentage of colonies onDox that exhibit morphological differences from their counterpart controls (Never state) was 30.2% in stem cells (Figure 26. C). This percentage difference was similar to that of Neu (28.5%). In contrast, almost all of the luminal progenitor cells (98%) responded to the activation of Myc and grew into filled colonies. In case of Neu, there were no apparent morphological and proliferative differences between the luminal progenitors onDox and the never induced controls. Again, the Myo hi population gave rise to aberrant colonies (32.3%) that are morphologically similar to the stem cell derived ones. Despite sharing the same MMTV-promoter, we again observed differences in the expression pattern of the transgenes between Neu and Myc.

Interestingly, in line with the previous report by Joshi *et al.*, depending on the phase of estrus cycle when the mammary glands were isolated, the total number of colonies arising in culture differed significantly[98]. However, the percentages of colonies responding to induction of oncogene (i.e. giving rise to aberrant morphological changes) were fairly constant between steady state (estrus) and highly proliferating (diestrus) state (data not shown) upon both Neu and Myc overexpression.

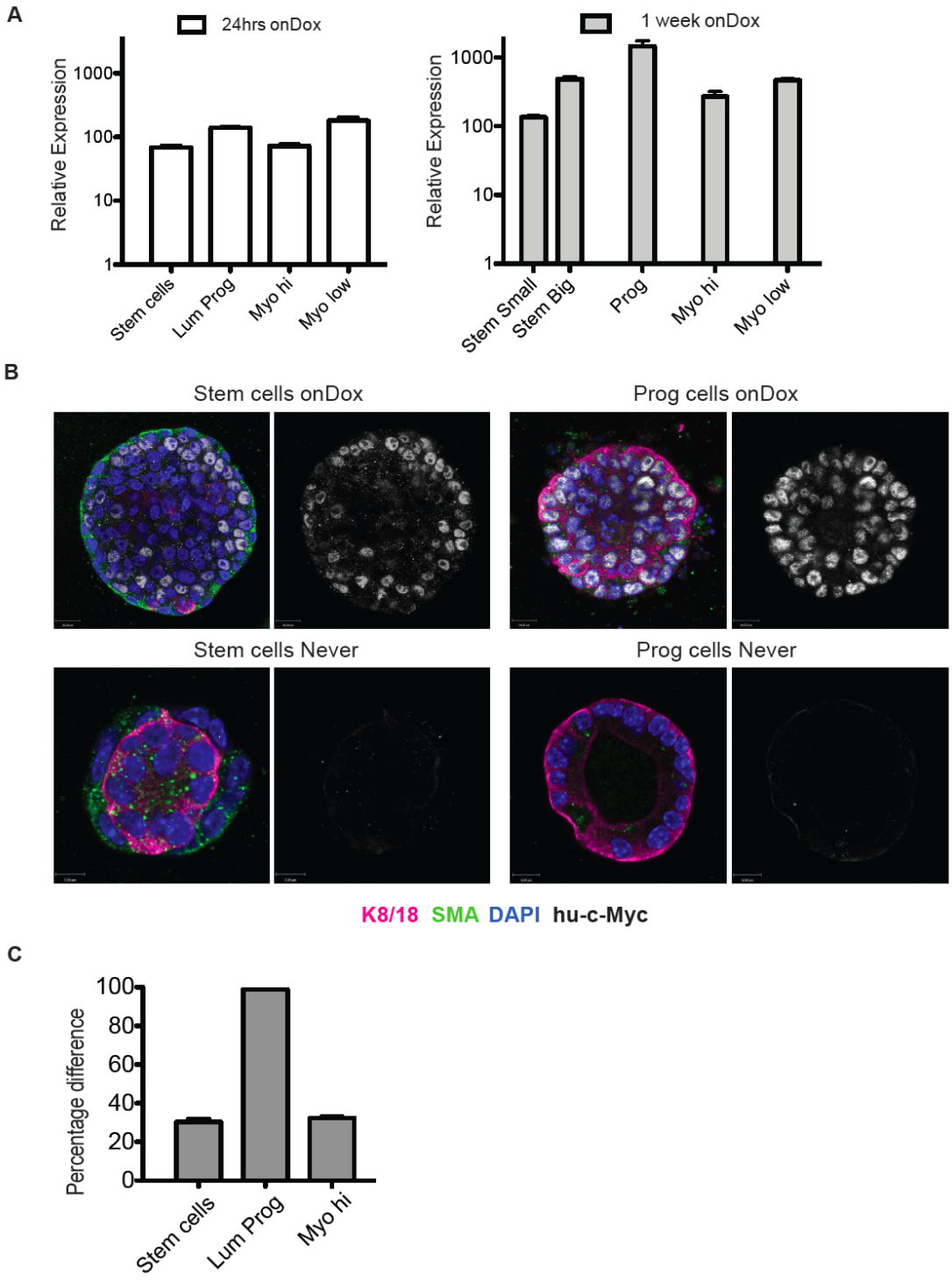


Figure 26. *in vitro* culture recapitulate early stage of Myc driven tumorigenesis *in vivo*. A) Relative expression of human-c-Myc transgene in mRNA levels in the different FACS enriched cell populations detected by qRT-PCR. Y axis showed relative level of mRNA expression in logarithmic scale. Different fractions of FACS enriched cell population from T-O-MYC express the transgene hu-c-Myc in culture. B) Immuno-staining of the transgene hu-c-Myc expression in cultured colonies (1 week in culture). C) A fraction (30.3%) of induced stem cells and the majority (98.9%) of progenitor cell colonies in culture exhibit significant morphological differences upon doxycycline induction compared to the never induced controls.

The bright field analysis of the colonies in culture showed that after 7 days, both stem and progenitor colonies start to have some degree of cell death. In case of the stem cell colonies overexpressing Myc, the inner portion of the colonies has dead cells (Figure 27. A). On the other hand, luminal progenitor colonies in onDox culture seem to have cell death from the outside (Figure 27. B). Myo hi population presented similar features to that of stem cells, but Myo low population gave no differences in culture (Figure 27. C and D). The effect of oncogenic Myc was rapid and potent at 1mg/ml of doxycycline administration in culture. The activation of apoptotic, proliferative pathways by Myc overexpression resulted in both expansion and cell death of colonies in culture. For this reason, the FACS enriched populations from T-O-MYC were kept in culture no longer than 2 weeks. Alternatively, a lower concentration of doxycycline can be used in culture in the future.

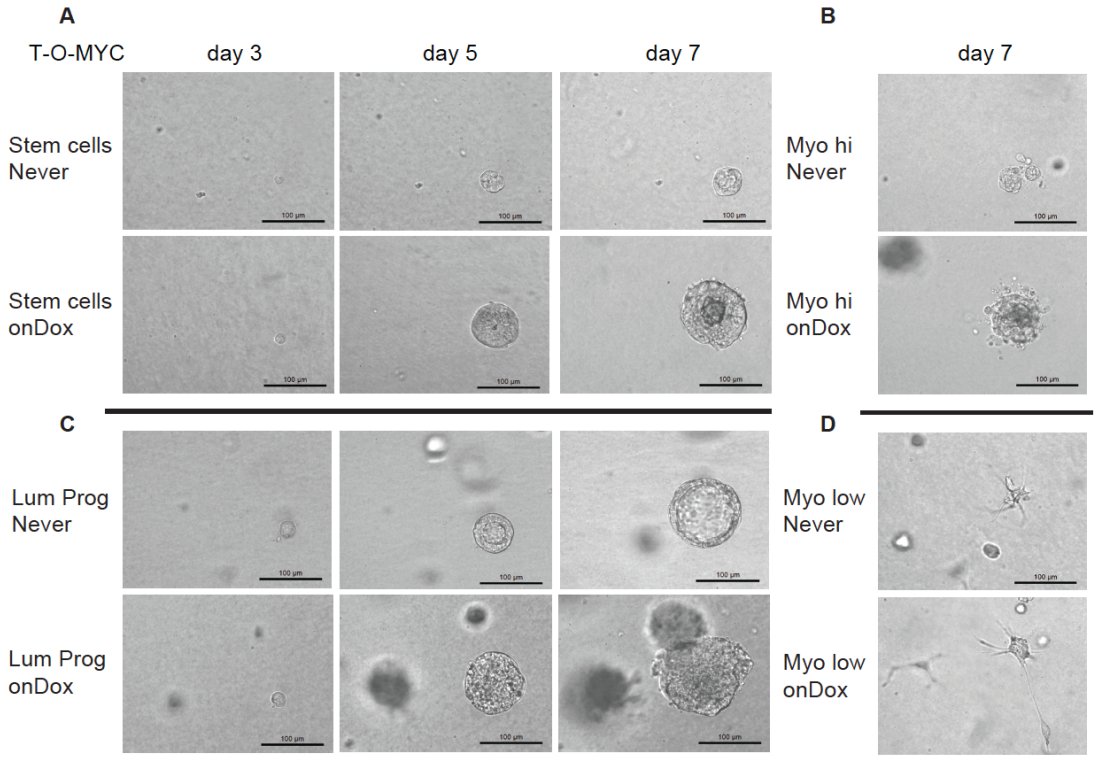


Figure 27. Bright field photo documentation analysis of the sorted cells from T-O-MYC in culture over 1 week of period. A) Representative images of stem cell colonies. A fraction of the induced stem cells (onDox) responds to transgene activation. B) Representative images of Myo hi colonies. Similar response was observed in Myo hi population. C) Representative images of luminal progenitor acini. Nearly every colonies give rise to morphologically aberrant colonies in response to Myc overexpression. D) Representative images of Myo low population. No significant difference was seen in Myo low population.

We were able to confirm the increase in proliferation of Myc overexpressing stem and luminal progenitor colonies through incorporation of EdU in culture (Figure 28. A-

B). Upon Myc oncogene overexpression, the expressions of lineage specific markers (SMA and K8/18) were analyzed. Among the bigger stem cell derived colonies (became over 100μm in diameter), the number of cells express K8/18 luminal lineage decreased (Figure 28. C). Majority of stem cell derived colony kept the expression of SMA especially at the outer layer of colonies. There were some cells in the middle of stem cell colonies that were negative for both SMA and K8/18 (e.g. day 10). We assessed if it was due to lack of antibody penetration; however, that was not the case.

The luminal progenitor cells from T-O-MYC showed continuous expression of luminal lineage marker K8/18, but Myc overexpression also increased the level of SMA expression in those colonies (onDox). Elevated SMA expression among onDox luminal progenitor colonies resulted in the double expression of luminal and myoepithelial lineages in early time points (day 5). Over time, the K8/18 expression became heterogeneous within the luminal progenitor colonies. Especially, the cells located at the outer layer showed a reduced K8/18 expression (at day 7). Furthermore, at day 10 of culture, onDox colonies had some single SMA positive cells in the outer layer of colonies that no longer express K8/18. Interestingly, onDox cultured progenitor colonies resembled more of the cellular composition of these Myc primary tumor, i.e., they contained both luminal and myoepithelial cellular lineages. While we cannot exclude the possibility of the stem cells being the TICs, these *in vitro* observations suggest that luminal progenitor cells could also be the TICs of Myc driven tumors.

In contrast to Neu, overexpression of Myc in luminal progenitors resulted in the preferential expansion of the myoepithelial lineage that could play a role during Myc induced tumorigenesis. We further analyzed this phenomenon of cellular transition in different experimental settings both *in vitro* and *in vivo*.

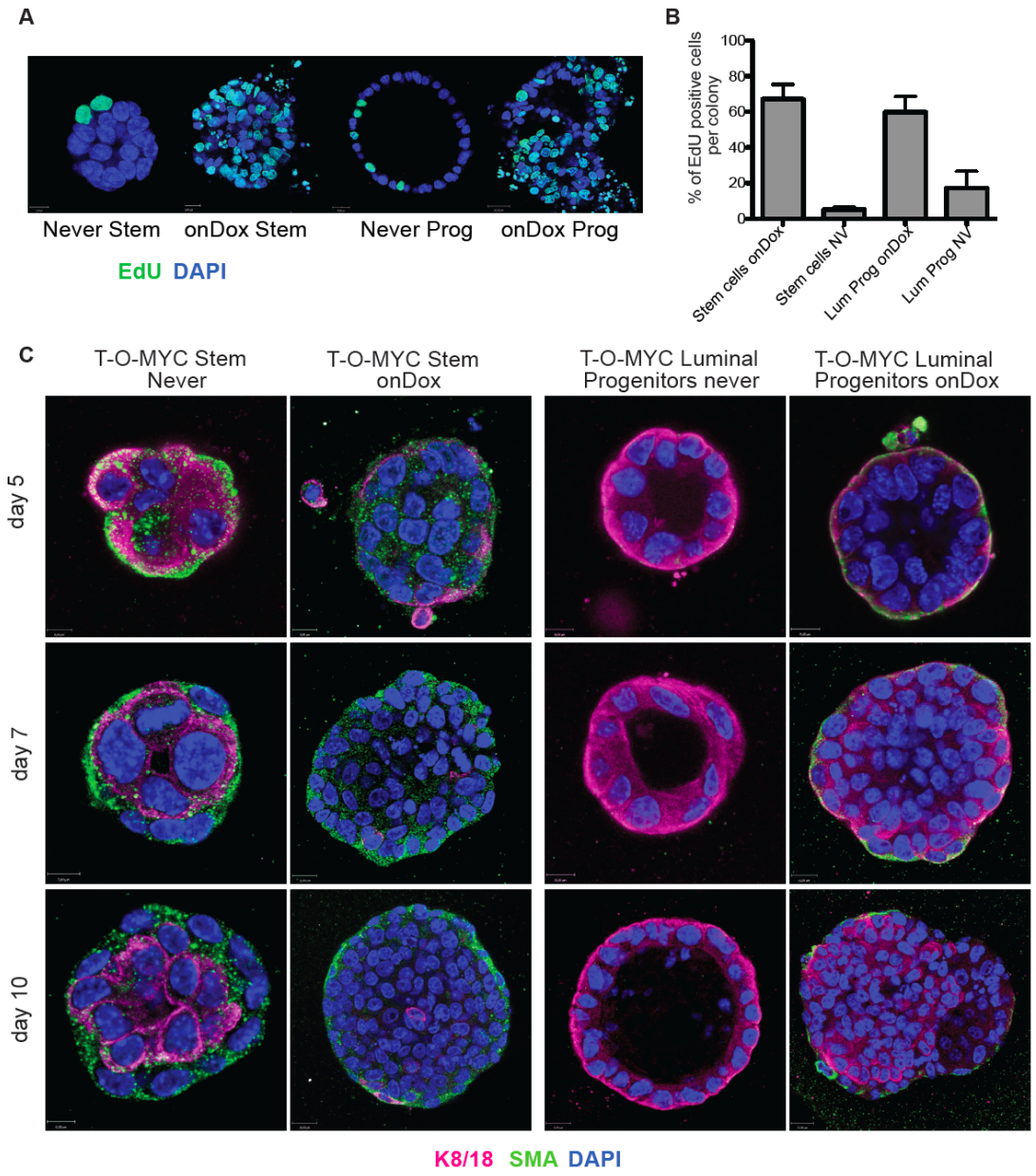


Figure 28. A fraction of stem cells and majority of luminal progenitors responded to Myc overexpression in culture by increasing their proliferating and preferential differentiation towards myoepithelial lineage. A) Representative pictures of Stem and progenitor cells with and without doxycycline induction for 3 weeks, showing an increase number of positive EdU cells on the SC and Lum Prog onDox. Green: EdU staining, blue: DAPI (EdU incorporation 18 hours) B) Quantification of EdU incorporation in the spheres in culture (n>10 colonies). C) Immuno-fluorescence staining of cellular lineages (luminal and myoepithelial) of the cultured colonies. Stem cells colonies shows loss of luminal lineage and expansion of SMA positive myoepithelial lineage. Luminal progenitors also gain the expression of myoepithelial lineage upon doxycycline induction. Over a period of one week in culture, the luminal progenitors express single positive myoepithelial lineage in the outer layer of the colonies.

3.7 Slug up-regulation indicate myoepithelial lineage commitment in Myc and Neu driven colonies *in vitro*

Previous reports suggested that an elevated level of Slug, a critical transcription repressor of breast epithelial differentiation, can be detected within the myoepithelial lineage[158,159]. We first attempted to confirm the expression of Slug in wild type or the Never induced cultured colonies. As expected, in luminal progenitor population, there was no Slug expression detected by immuno-fluorescence staining. Among the stem cells colonies, Slug was up-regulated only in the colonies where the myoepithelial/basal cellular differentiation took place usually after 7 days of culture (Figure 29. A). When the stem cells were not differentiated into a bi-layer of luminal and myoepithelial acini, Slug was not expressed. Upon Myc overexpression, the stem cell colonies expand their myoepithelial lineage shown by SMA expression. As expected, we confirmed an up-regulation of Slug within the nuclei of SMA positive cells (Figure 29. B). This also suggests a preferential expansion of the stem cells into the myoepithelial-committed lineage. Similarly, we attempted to show the lineage transition of luminal progenitor colonies upon Myc overexpression. Potential trans-differentiation from luminal to the myoepithelial lineage can be confirmed by Slug staining together with SMA, K8/18 expression. The never induced luminal progenitors were clearly Slug negative (Figure 29. C). In favor of our hypothesis, Slug up-regulation was also found within the single-SMA positive cells in the outer layer of the onDox luminal progenitor colonies (Figure 29. D, day 10). In addition to the cellular lineage marker SMA and K8/18, the co-immuno-staining of Slug suggests a trans-differentiation of luminal progenitors to the myoepithelial lineage upon Myc overexpression after a given period of time.

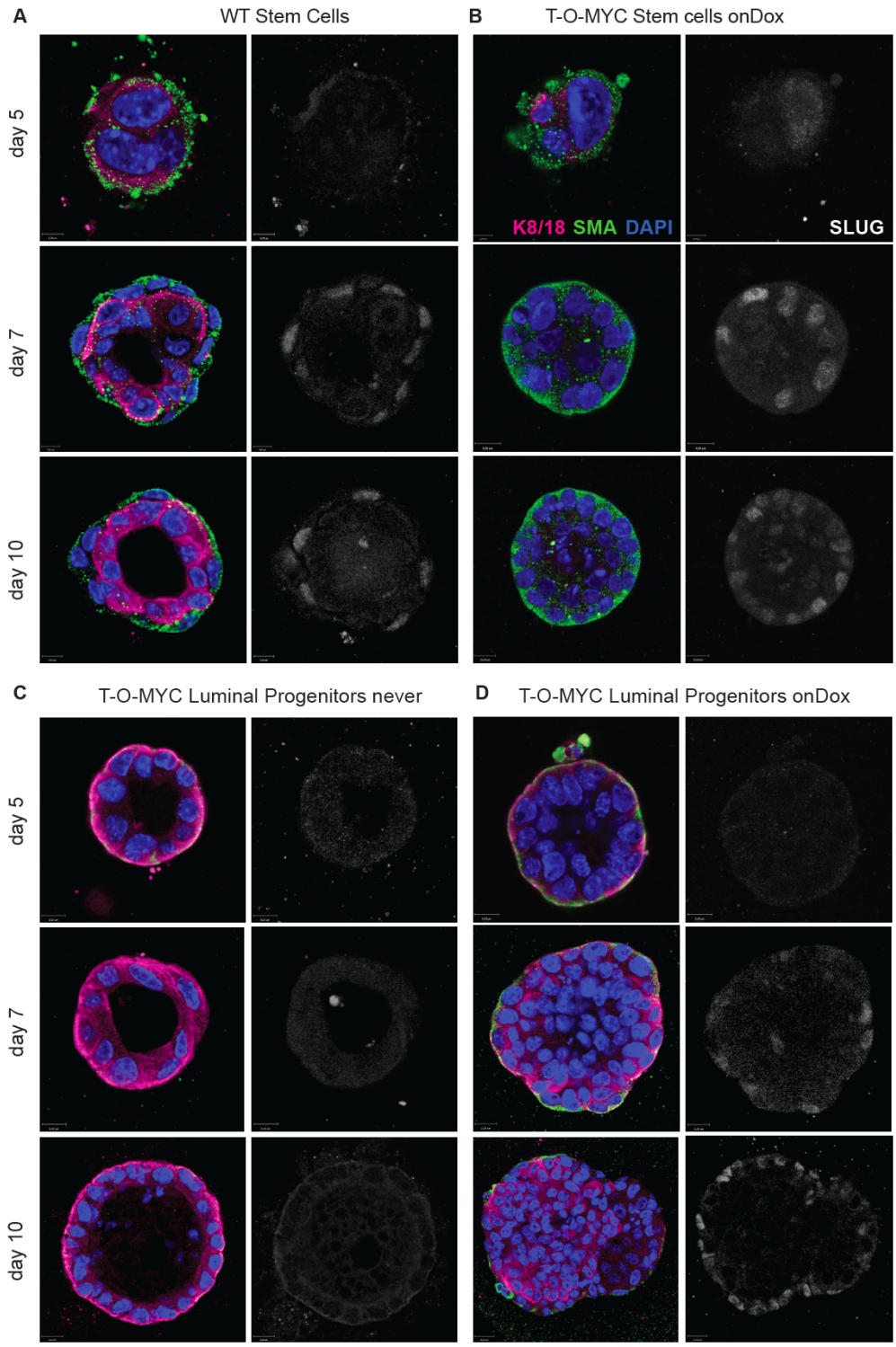
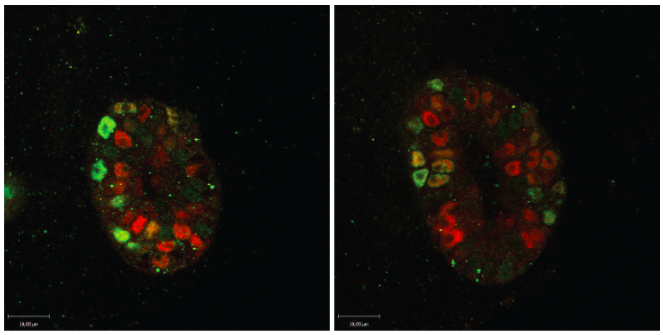


Figure 29. Slug up-regulation indicates myoepithelial lineage commitment in Myc driven colonies *in vitro*. A) Immunofluorescence staining of Slug, SMA and K8/18 in the wild type stem cell colonies in culture at different time points. Slug is up-regulated where the stem cells differentiate and commit to the myoepithelial lineage. B) Immunofluorescence staining of Slug, SMA and K8/18 in the T-O-MYC stem cell colonies in culture onDox at different time points. The outer layer of SMA positive cells frequently co-expressed Slug. C) Immunofluorescence staining of Slug, SMA and K8/18 in the Never induced luminal progenitors where Slug is not expressed. D) Immunofluorescence staining of Slug, SMA and K8/18 in the T-O-MYC luminal progenitor colonies onDox at different time points. Single SMA positive cells within the luminal progenitor cells onDox are up-regulated in the level of Slug expression.

Since the majority of Slug positive cells are located in the outer rim of the acini, we confirmed again the expression of transgene hu-c-Myc with immuno-staining. Within the luminal progenitor colonies from T-O-MYC, we confirmed a clear Slug up-regulation in the outer layer of the colonies (Figure 30). Myc transgene expression was not limited to the outer rim of the colonies.



hu-c-Myc Slug

Figure 30. Representative images of immuno-staining of hu-c-Myc and Slug expression among T-O-MYC luminal progenitors after 7 days onDox.

3.8 Slug expression status in Neu overexpression

In contrast to Myc, Neu overexpression within the stem cell colonies led to down-regulation of SMA expression (Figure 25. C). As it is suggested that Slug down-regulation correlates with the luminal lineage expansion[158], we attempted to confirm the preferential expansion of luminal lineage. The K8/18 expressing aberrant stem cell colonies from T-O-NEU exhibit no Slug up-regulation in the inner part, however, we found the mild Slug up-regulation in the outer layer of the colonies (Figure 31. A). This implies that the cells in the outer layer of Neu overexpressing stem cell colonies lose the SMA expression, but still keep their myoepithelial lineage. We found those SMA negative cells still express the un-committed myoepithelial marker (Figure 31. B) We also questioned that if the loss of SMA positive cells within the Neu overexpressing stem cell colonies is due to programed cell death. However, there were no sign of Caspase-3 dependent apoptosis in the outer layer of stem cell colonies overexpressing Neu (Figure 31.C). We have not confirmed whether longer overexpression of Neu would eventually down-regulate the Slug expression completely.

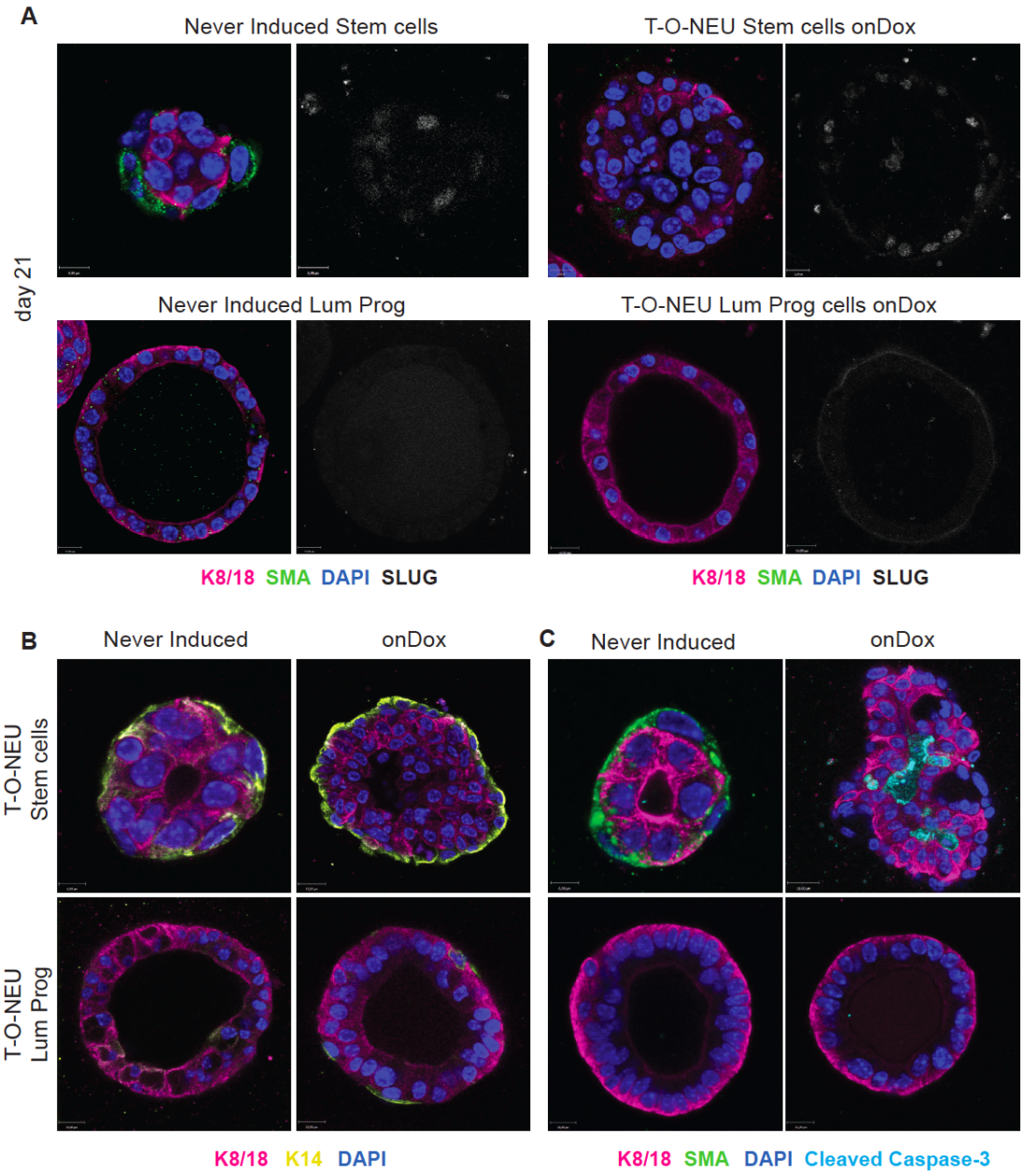


Figure 31. The outer layer of cells within the Neu overexpressing stem cells are still shown to be un-committed myoepithelial lineage. A) Representative images of Slug, SMA and K8/18 expression within the stem cells and luminal progenitor cells from T-O-NEU. B) Representative images of K14 and K8/18 staining within the stem cells and luminal progenitor cells from T-O-NEU. C) Representative images of SMA, K8/18 and cleaved Caspase-3 staining within the stem cells and luminal progenitor cells from T-O-NEU.

3. 9 Neu target the un-committed stem cells as TICs *in vitro*

Despite the expression of Neu transgene in the progenitor population, *in vitro* cultured cells onDox did not show any sign of tumorigenicity. This puzzled us since previously, the luminal progenitors were thought to be the targeted population or the

origin of TICs of Neu driven tumors. We attempted to then test different micro-environmental factors that could be lacking in our culture condition. First, we included fat cells (adipocytes) in the co-culture. This increased the proliferation of progenitor cells in general, but the differences were not significant between onDox and never induced controls (data not shown). Next, we examined the effect of myoepithelial cells surrounding the luminal cells. We attempted to examine this by first, letting T-O-NEU stem cell population expand in culture for 2 weeks without doxycycline induction, in order to differentiate into myoepithelial- and luminal- cells containing bi-layered colonies. Stem cells gave rise to a heterogeneous population including three main categories; 1) small slow-proliferating stem-like colonies (SMA and K8/18 double positive), 2) partially differentiated bi-layered colonies, and 3) luminal progenitor like single layer acini (Figure 32 Upper). Although these results are preliminary, we found that Neu overexpression led to an increased proliferation in the stem like colonies (category 1) and some bi-layered colonies (category 2), but rarely in the luminal progenitor-like acini (category 3, <5%) (Figure 32 Lower). This implies that cellular differentiation status plays an important role for the cells to be more susceptible to the hit of the oncogenic mutation and become TICs. In order to confirm this result, further lineage tracing approaches must be employed.

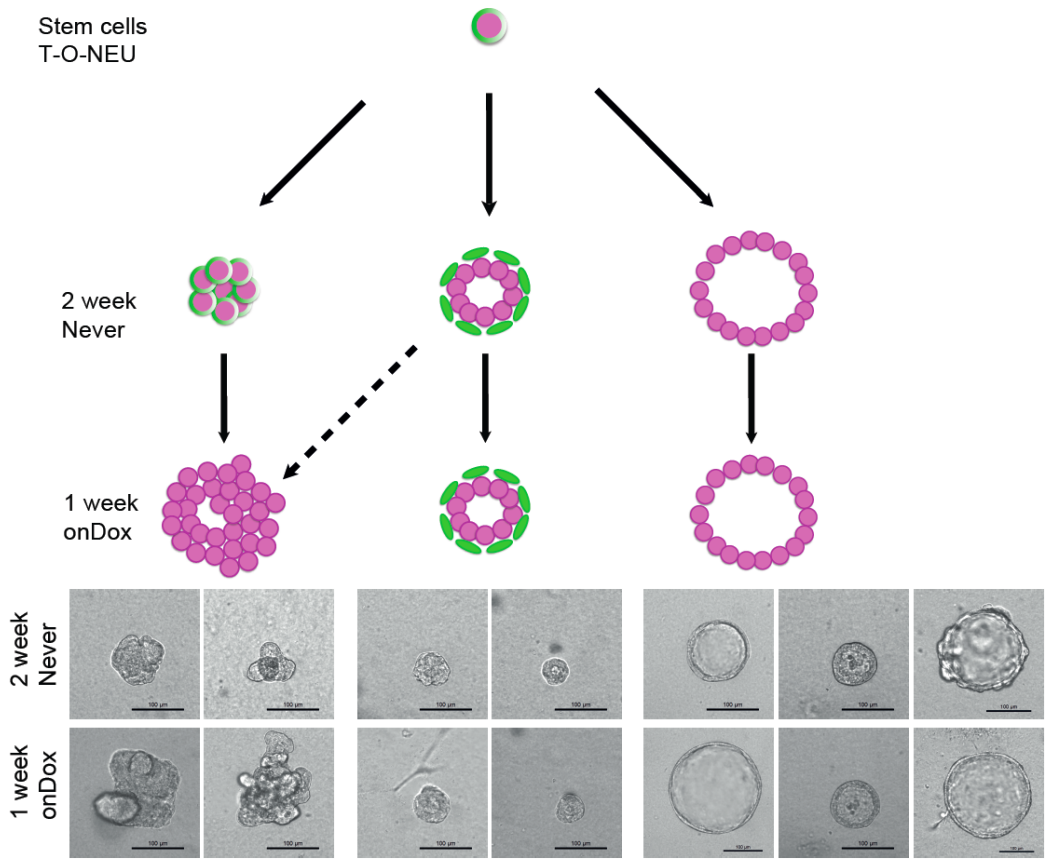


Figure 32. Neu target the un-committed stem cells as TICs *in vitro.* Upper) Schematic representation of three different cellular categories derived from stem cells and their response to Neu overexpression. Lower) Representative bright field images of T-O-NEU stem cells derived colonies.

3.10 *in vivo* reconstitution of mammary gland for the tumorigenesis assay

In parallel with the *in vitro* culture approach, we employed *in vivo* transplantation assays in order to assess the tumorigenic capacity of potential TICs. As previously described, we enriched and isolated stem and luminal progenitor cells by FACS from the digested mammary glands of un-induced T-O-NEU and T-O-MYC 8 weeks old adult female mice. The FACS enriched cells were re-suspended in 50% of matrigel (BME) containing media, and transplanted into immune-compromised $Rag1^{-/-}$ mice. This transplantation could provide one of the closest microenvironment for the FACS isolated mammary epithelial cells to their original microenvironment. We injected the different populations of FACS isolated cells in serial dilutions obtained from un-induced bi-transgenic mouse mammary glands. Knowing the MRU of the stem cells (1/196) and luminal progenitors (1/1614) we decided to perform these transplantations in serial dilutions of 250 cells, 500 cells, 1000 cells and 2000 cells. The range should cover the cellular dilution when stem cells are likely to reconstitute the mammary fat pad upon transplantation while the luminal progenitor cells would not. In order to induce oncogene expression, the immune-deficient recipient mice were fed with doxycycline containing food immediately after the transplantation (Figure 7 in introduction). After 6-8 months, tumors developed from the transplanted cells injected in the cleared fat pad upon continuous administration of doxycycline (Figure 33. A and Figure 35. A). Control mice that did not receive doxycycline-containing diet did not develop any tumor upon transplantation. The cellular compositions of tumors obtained upon transplantation were then compared to the respective primary tumors (Figure 33. B and Figure 35. C).

3.11 Luminal progenitor cells as well as stem cells are the potential TICs in Neu overexpressing tumors *in vivo*

As previously reported, the luminal progenitor cells had been described as the TIC population of Neu/HER2 driven tumors shown in MMTV-NEU mice. From our *in vitro* observation, we suspected that stem cells could also be the TICs. Upon activation of Neu overexpression, we observed the formation of tumors from the stem, luminal progenitor and Myo hi population, but not efficiently from the Myo low stromal

population (Figure 33. A). The efficiency of tumor forming capacity, tumor initiating unit (TIU) was also calculated using ELDA. The stem cells were able to form tumors with an efficiency of TIU 1/530, whereas luminal progenitor cells also gave rise to tumors with a frequency of TIU 1/606. Interestingly, the Myo hi population could also give rise to tumors *in vivo*. In fact, we repeatedly observed that the Myo hi population presented similar characteristics to that of stem cells both *in vitro* and *in vivo*. Currently, there are no specific cell surface markers to distinguish Myoepithelial progenitor or Myoepithelial differentiated cells from the stem cells that also resides in the myoepithelial population. Although it is a mixed population, it is highly likely that the Myo hi population also contain cells that possess certain stem-like characteristics. In the *in vitro* culture, we did not observe any significant tumorigenic characteristics of the luminal progenitor cells, however, luminal progenitor cells gave rise to tumors as efficiently as that of stem cells.

The immuno-staining with SMA and K8/18 confirmed that the tumors from transplanted cells were mainly composed of luminal cells in all cases (Figure 32. B Upper). This resembles the classical Neu driven luminal tumor phenotype as well as the primary tumors obtained from T-O-NEU (Figure 11. A). We could also confirm the HER2 expression at the protein level (immunostaining) among the tumors arising upon transplantation (Figure 32. B Lower). Altogether, these results suggest that both stem and progenitor cells are the source of TICs in Neu driven tumors.

In order to confirm the contribution of T-O-NEU transplanted cells for their tumorigenesis, we attempted to mark the tumor cells originating from the transplanted cells. We crossed T-O-NEU mice with CAG-H2B-GFP transgenic line to be able to distinguish the transplanted cells from the cells of the recipient mice. Surprisingly, we observed down-regulated GFP expression in some tumors that originated from stem cells (n=3) (Figure 34. A). Interestingly, the neighboring mammary duct still expressed GFP in the nuclei only among the myoepithelial cells and not the luminal cells. Some luminal progenitor derived tumors expressed GFP (2 out of 4) (Figure 34. B), but only in some areas (<10%). Although this is a preliminary analysis, GFP expression is either down-regulated, silenced or selected against upon Neu overexpression and tumor initiation.

A

T-O-NEU	Number of cells injected per fat pad	Number of tumor outgrowths	TIU
Stem cells	2000 1000 500 250 100 50	15/17 4/5 7/8 0/1 2/3 0/2	1/530
Luminal Progenitors	2000 1000 500 250 100 50	14/15 4/4 6/10 0/4 0/4 0/2	1/606
Myo hi	2000 1000 500 250 100 50	3/5 - 3/4 - - -	1/1237
Myo low	2000 1000 500 250 100 50	2/3 - 0/3 - - -	1/2624

B

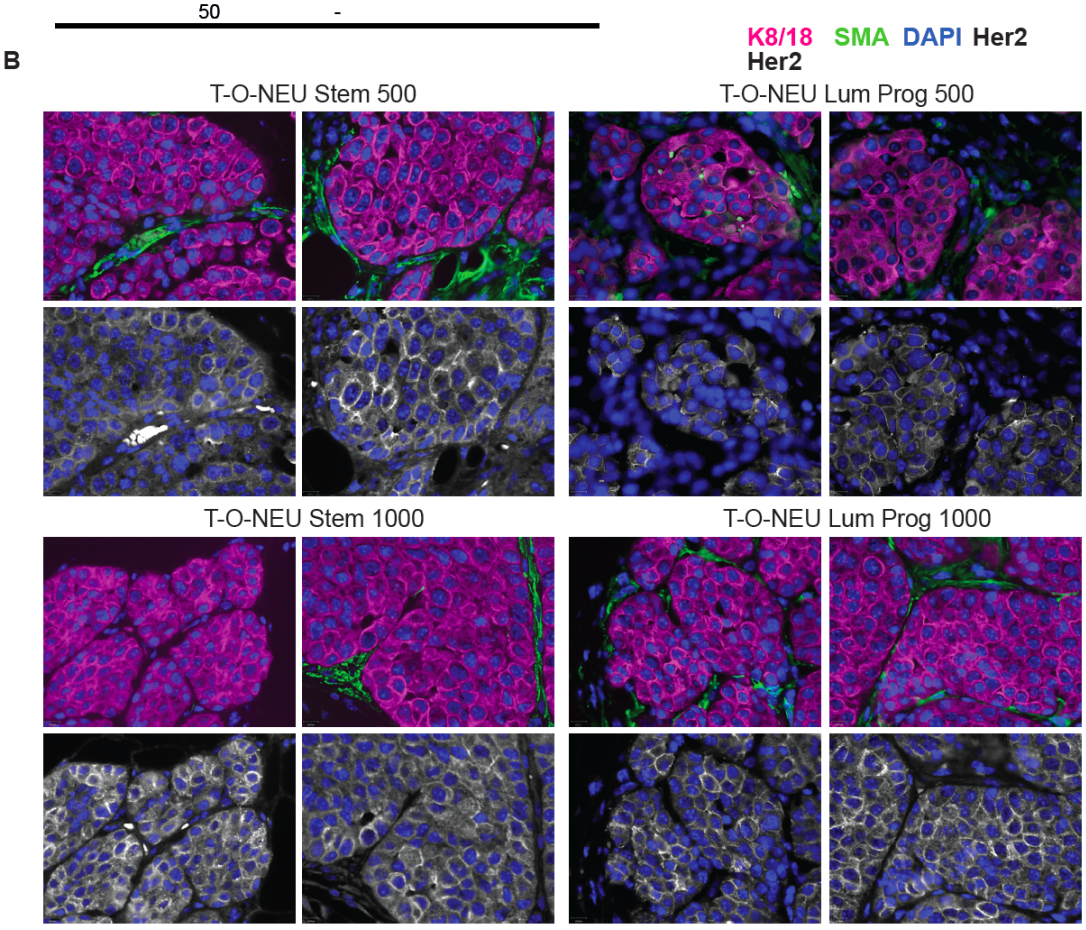


Figure 33. Luminal progenitor cells as well as stem cells are the potential TICs in Neu overexpressing tumors *in vivo.* A) Table summarizing the result of *in vivo* tumorigenesis assay and estimation of tumorigenic efficiency of T-O-NEU sorted cells *in vivo.* B) Panel of immuno-staining of tumors arising from the transplantation of T-O-NEU sorted cells (upper), confirming HER2 transgene expression (lower).

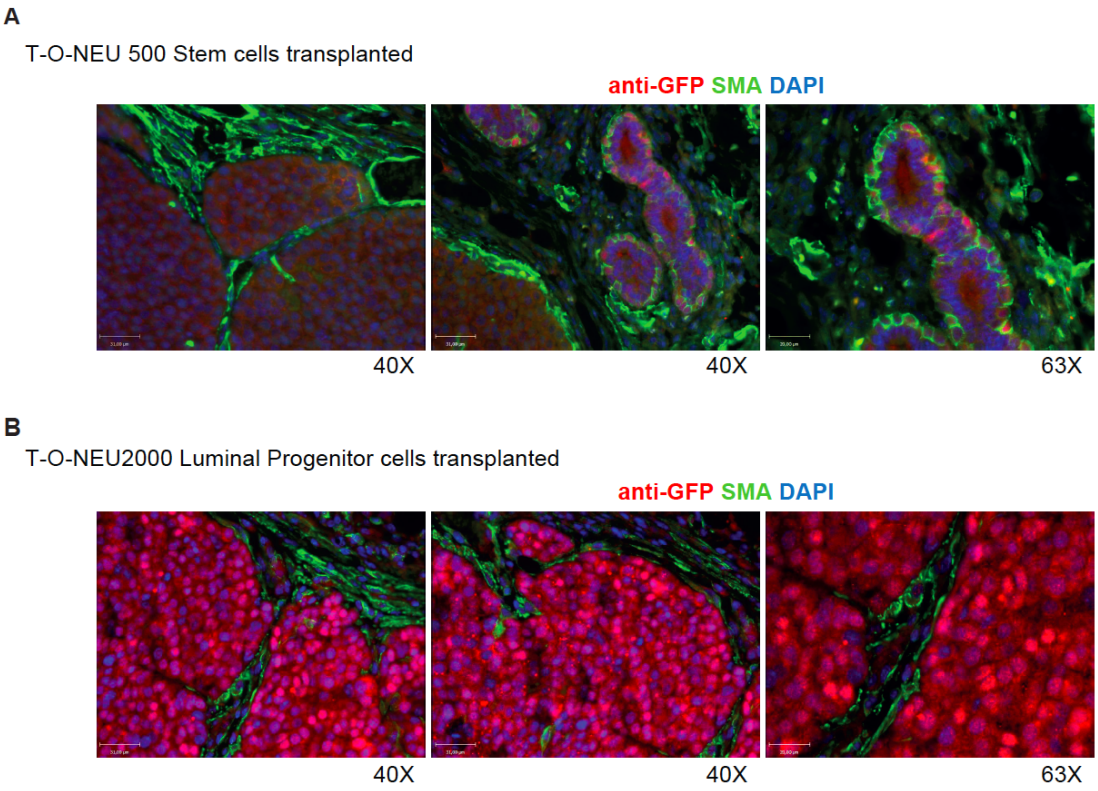


Figure 34. H2B-GFP is not always expressed in the tumors derived from transplantation of T-O-NEU/H2B-GFP cells. A) Representative images of tumors developed upon transplantation of stem, and B) luminal progenitors (lower). Myoepithelail cells are H2-GFP negative in both cases.

3.12 The potential TICs in Myc overexpressing tumors *in vivo*

TICs of Myc induced tumors were thought to be the stem cell population that give rise to heterogeneous tumors of a mixed phenotype. Upon transplantation of different FACS enriched T-O-MYC cell populations, we indeed obtained tumors from the stem and Myo hi lineage both composing the myoepithelial population[129]. Repeatedly, Myo hi population showed similar results to that of stem cells. We also observed some tumors from the luminal progenitor population but in a lesser frequency (Figure 35. A). Due to the relatively long tumor latency, also in the case of primary T-O-MYC tumors, the results are still preliminary.

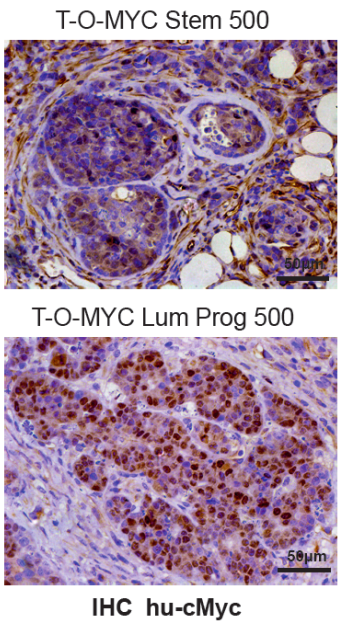
We confirmed human-c-Myc expression among the tumors upon transplantations (Figure 35. B). Analysis of the cellular lineage markers (SMA, K8/18) showed that all of those tumors are composed of both luminal and myoepithelial lineage (Figure 35. C). This resembles the primary T-O-MYC derived tumors (Figure 11. D). T-O-MYC/H2B-GFP tracing experiment was also performed in order to confirm the contribution of transplanted cells for tumorigenesis. The approach could confirm the

trans-differentiation of luminal progenitors to myoepithelial lineage upon Myc overexpression previously observed *in vitro*. Unfortunately, due to the long latency of Myc induced tumors (more than 6 months in average), the corresponding results cannot be shown in this thesis and will be presented somewhere else.

A

T-O-MYC	Number of cells injected per fat pad	Number of tumor outgrowths	TIU
Stem cells	2000 1000 500 250 100 50	1/4 0/3 1/4 0/3 0/1 0/1	1/4180
Luminal Progenitors	2000 1000 500 250 100 50	1/4 0/1 1/4 0/3 0/1 0/1	1/6297
Myo hi	2000 1000 500 250 100 50	2/3 0/2 1/3 1/2 - -	1/1813
Myo low	2000 1000 500 250 100 50	0/2 0/2 0/2 0/2 - -	n/a

B



C

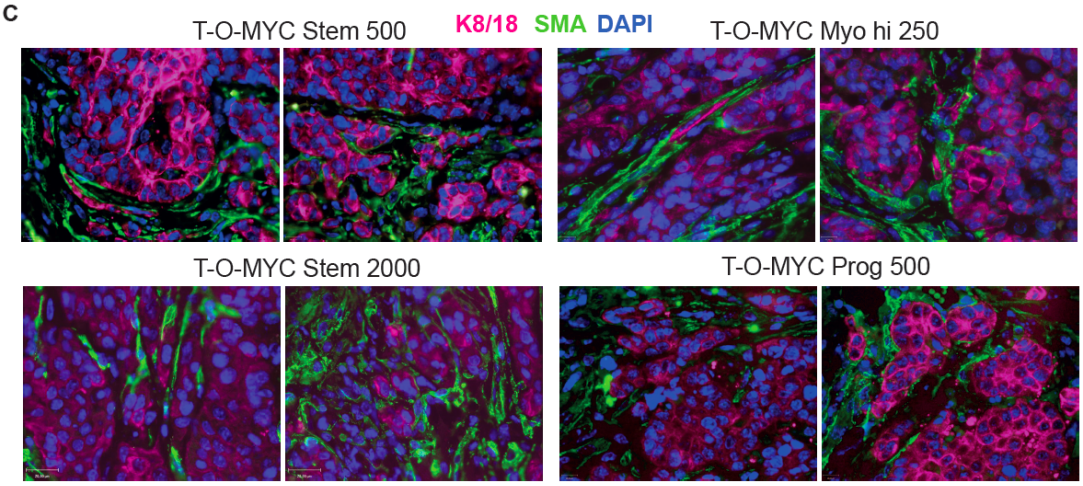


Figure 35. Stem cells as well as luminal progenitor cells are the potential TICs in Myc overexpressing tumors. A) Table summarizing the result of *in vivo* tumorigenesis assay and estimation of tumorigenic efficiency of T-O-MYC sorted cells *in vivo.* B) Panel of human-c-Myc immunohistochemistry of tumors arising from transplantation of T-O-MYC sorted cells. C) Immuno-staining of SMA, K8/18 in tumors obtained from transplantation.

3.13 Lineage tracing approach

In an effort to better characterize the TICs of Neu and Myc induced breast tumors, we employed two lines of transgenic mice which enable lineage tracing: MMTV-CreERT2 and Axin2-CreERT2. Axin2 has been shown as the stem cell marker at the adult stage (8 weeks) recently characterized by van Amerongen *et al.*[131]. We crossed these lines with two reporter lines Rosa26-YFP and Rosa26-Brainbow2.1 (Confetti).
First, in order to optimize the marking efficiency and to reproduce the published result, we induced Axin2-CreERT2/Rosa26-YFP at 8 weeks of age. After one week, and 4 weeks of tracing interval, we analyzed the YFP (Yellow Fluorescent Protein) positive cells using a FACS approach. At one week, a small portion of Myo hi population and 10% of Myo low population were YFP positive (Figure 36. A). After another 3 weeks of tracing, YFP positive cells were found in 7% of stem cells, 7% of Myo hi (Figure 36. B). The Myo low YFP positive cells remained constant (10%). These results indicate that indeed Axin2 is able to mark putative stem cells, however, induction of tamoxifen at 8 weeks also label some Myo low cells. Myo low population is known as the stromal cells and Axin 2 positive Myo low YFP positive cells did not show any expansion between 1 week and 3 weeks of tracing time from the FACS analysis. There is a possibility that some YFP positive Myo low population could give a false negative or a false positive results in the future tumorigenesis tracing experiment, however, it is less likely that Myo low population serves as the TICs upon overexpression of oncogenes. This is also supported by our previous *in vitro* and *in vivo* experiments. In parallel, MMTV-CreERT2/Rosa26-YFP lines were also characterized (data not shown). After the initial validation, we started to induce Neu and Myc oncogenes and assess the contribution of marked cell populations for the formation of tumors by further crossing with T-O-NEU or T-O-MYC mouse models (Table 7). Due to the time limitation, we cannot present these results in this thesis. However, this approach will be continued further and presented elsewhere.

Table 7. Various mouse models that are used for lineage tracing during tumorigenesis.

Methods of Analysis	**Mouse lines**
FACS	Axin2-CreERT2/Rosa26-YFP/TetO-Neu/MMTV-rtTA
FACS	Axin2-CreERT2/Rosa26-YFP/TetO-Myc/MMTV-rtTA
Histology	Axin2-CreERT2/Rosa26-Brainbow2.1/TetO-Neu/MMTV-rtTA
Histology	Axin2-CreERT2/Rosa26-Brainbow2.1/TetO-Myc/MMTV-rtTA

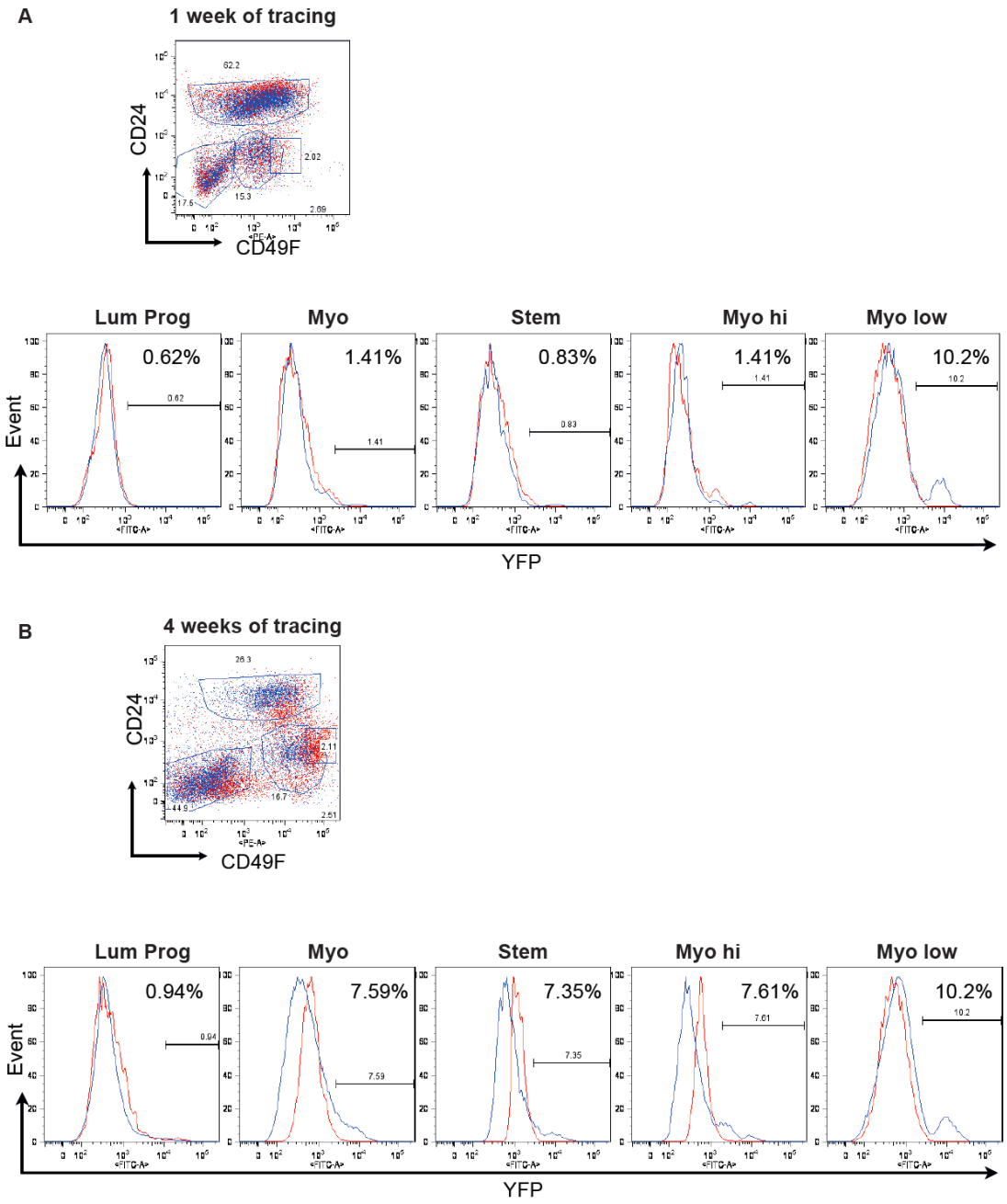


Figure 36. Initial FACS analysis of showed Axin2-YFP positive cells expanded in the Stem, and Myo hi populations over 4 weeks of lineage tracing. A) 1 week of tracing. B) 4 weeks of tracing. The experimental animals are Axin2-CreERT2/Rosa26-YFP (in blue) and the respective controls are tamoxifen-induced littermates Rosa26-YFP (in red). In order to analyze the percentage differences between the samples and their controls, the baseline was set as the 1% of YFP^{hi} populations of the controls. (Data jointly produced with Marta Garcia)

4. Discussion

4.1 Constitutive versus tractable mouse mammary tumor models

The majority of previous reports indicated that the TICs of MMTV-Neu induced tumors are the luminal progenitors[89,90,160]. All these studies were conducted using MMTV-Neu, and the effect of Neu overexpression were present during embryonic and pubertal development in all mammary cellular lineages in contrast to the *de novo* activation of Neu in an adult cell. The constitutive overexpression of Neu may also alter the physiological mammary cellular hierarchy during development. Unlike mutations of the tumor suppressor gene such as BRCA-1 that is classified as a hereditary or germ line mutation, Neu or c-Myc oncogenic mutations appeared to be somatic mutations in patients that are acquired after puberty and throughout adulthood. In order to delineate the TICs in adult healthy mammary gland more accurately; we employed tetracycline inducible mouse mammary tumor models that closely recapitulate somatic oncogenic mutations in different adult mouse epithelial lineages including stem cells, and luminal progenitor cells.

4.2 Use of *in vivo* transplantation assay to assess tumorigenicity

Self-renewal and lineage differentiation capacity are two main characteristics of all stem cells. Therefore, as with normal stem cells, evaluations of both self-renewal and tumor propagating capacity of TICs were critical. The transplantation assay and confirmation of tumor growth *in vivo* could fulfill these criteria. Although there are some existing withdrawals, *in vivo* transplantation is regarded as the gold standard functional assay.

In our study we performed, first, the orthotopic transplantation of healthy mammary epithelial cells into the cleared mammary fat pad. The mammary gland is one of the rare organs that can be dispensable and fully reconstituted upon transplantation. We did not use syngeneic mice that are classical recipients of mammary stem cell transplantation, but instead, $Rag1^{-/-}$ mice as the host. $Rag1^{-/-}$ mice lack B and T lymphocytes while the other immune cells including natural killer (NK) cells are still present[65]. Other common immuno-deficient mice used for transplantation assays are NOD/SCID mice or nude mice. In the wild type situation, upon transplantation of bi-potent stem cells, we were able to reconstitute functional mammary gland in cleared

mammary fat pad of Rag1$^{-/-}$ recipient mice. Our transplantation efficiency was 1/2-1/3 times lower compared to the previous reports[118,128]. This could be due to technical variability, genetic background of bi-transgenic mice, and the use of Rag1$^{-/-}$ mice as the hosts instead of using syngeneic mice. Hence the difference could be due to presence of NK cells and other immune cells still present in the recipient mice.

From our *in vivo* tumorigenesis assay, the luminal progenitor cells indeed gave rise to tumors that resemble the primary breast tumors from T-O-NEU. However, our putative stem cells and Myo hi population also could give rise to tumors that are composed of mainly the luminal cells similar to the primary tumors. Moreover, the efficiency of tumor initiation (shown in TIU) was slightly higher among stem cells than that of luminal progenitors. From *in vivo* data, it hints that adult stem cells and luminal progenitor cells are putative TICs of T-O-NEU mouse mammary tumor models. We did not perform serial transplantation of tumor cells obtained upon transplantation as our aim is not to identify the CSCs; identification of CSCs often involve serial transplantation assay.

The stem cell being a potential TIC of Neu driven tumors contradicts with the recent reports postulating that the luminal progenitors are the TICs[89-91]. The PI-MECs are also suggested to be the TICs of Neu driven tumors[87], however, we did not examine stem cells that are expanded during pregnancy (PI-MECs). This could be a possible future experiment.

Upon overexpression of c-Myc, stem cells, Myo hi population gave rise to tumors and luminal progenitors to a lesser extent. From the primary breast tumors from T-O-MYC the median tumor free latency was double the time of T-O-NEU. Current results suggest that the myoepithelial lineage including the stem cells and Myo hi populations are the potential TIC of Myc driven tumor. This is in line with our initial hypothesis as the primary Myc overexpressing tumors in human and mice exhibit basal/myoepithelial mixed phenotype. More conclusive result from *in vivo* assay will be presented upon collection of more data elsewhere.

4.3 Transplantation assay – artifact or effect of inducing de-differentiation partially committed progenitor cells

The *in vivo* transplantation assay is rather a long process. On top, we try to assess the tumorigenic capacity of different cell populations that are induced overexpression

of oncogenes. The time line of tumor development after transplantation can take over 6 months. The activation of transgenes was achieved soon after the transplantation (within the first 24 hours) within the time frame that the cells of interests are likely to maintain their lineage characteristics and cross examined by *in vitro* culture. However, we cannot exclude the possibility of differentiation, de-differentiation variability induced by the transplantation procedure by itself.

Not only the stem cells but also upon injection of luminal progenitor cells, we were able to reconstitute mammary glands, but only when a higher number of cells are injected into the fat pad. Their reconstituted mammary gland spread to a lesser extent (<30%) in the entire fat pad compare to the ones of stem cells (>80%). We tried to examine the functionality of reconstituted mammary glands by the expression of lineage specific markers including K8/18, SMA and K14[161]. We found that the reconstituted mammary glands derived from the luminal progenitor cells consists of bi-layered ducts similar to the control mammary glands. The presence of SMA positive myoepithelial cells found in the reconstituted mammary gland upon transplantation of luminal progenitors made us question their cellular origin. A recent lineage tracing study also pointed out that trans-differentiation is a potential outcome of transplantation assay that does not frequently occur in the physiological conditions[120]. Although it is infrequent, interchange of lineage or de-differentiation of lineage-committed stem/progenitor to become multi-potent in physiological condition can be exacerbated through transplantation procedures. It could explain why partially committed luminal progenitors can also give rise to the myoepithelial lineage upon transplantation show no particular defects in the reconstituted mammary gland.

4.4 *in vitro* culture recapitulate normal mammary gland development

Although the *in vivo* transplantation assay has been the gold standard for identifying stem cells and provides the indication for identifying the breast TIC populations, using animal models is time consuming and difficult for high-throughput quantification analysis. Moreover, we could not address how the individual transplanted cells of interest would respond to the respective oncogenes at the pre-neoplastic stage. Therefore, we attempted to develop a reliable *in vitro* assay that is quantitative, rapid, and less expensive but still sufficiently sensitive and is specific to the cells of interest[162]. Our *in vitro* culture system successfully allowed us to monitor individual cells for their proliferation and differentiation mechanism maintaining their cellular characteristics.

We first optimized an *in vitro* 3D culture system using wild type cells, where FACS enriched stem and other cell populations can be maintained and expanded. In our validation, we observed the stem cell being less proliferative than the luminal progenitors. Luminal cells usually developed into single layer hollow polarized acini that express luminal specific marker K8/18 dominantly. Our culture allowed the cells to proliferate for more than 4 weeks and can also be passaged. Among luminal progenitor acini in culture at 1 week, there was a mind or less dominant SMA expression in different degrees. $CD61^{hi}$ expressing luminal progenitor population showed a mind SMA co-expression. In contrast, larger portion of $CD61^{low}$ luminal cells give rise to SMA negative K8/18 positive acini. However, after 2 weeks in culture the difference between these two populations became insignificant implying the differentiation of $CD61^{hi}$ population to become more committed luminal lineage.

The stem cells maintained their bi-potent lineage in culture shown by stem cell colonies often being dominant double positive for both luminal (K8/18) and the myoepithelial specific markers (SMA). After over 1 week in culture, some stem cells started to differentiate into bi-layered acini having inner layer of luminal cells and outer layer of myoepithelial cells. Other cell populations obtained from FACS enrichment including Myo low were also kept viable in culture but did not display any clonal expansion.

It is obviously important for the validity of this 3D organotypic approach to demonstrate that the respective cells would still behave like the original sorted population. We followed this, isolated cells and could show that stem cells in culture for 1 week were still able to give rise to the functional bi-layered mammary glands upon transplantation. Stem cell colonies in culture for over 1 week started to develop a degree of heterogeneity; some cells are staining as small colonies that have no lumen formation, some others starting to partially differentiate into the bi-layered polarized acini. Shorter period in culture could increase their MRU, the reconstitution efficiency.

Luminal progenitor cells maintained in culture were also able reconstitute the mammary gland, but with a lower efficiency. This could be due to the effect or artifact of transplantation assay as reported previously[120]. There is a small possibility that our FACS enriched stem cell populations are contaminated with other cell populations such as luminal progenitors. However this is unlikely, given the high FACS sorting

efficiency (over 96%) and the stem cell population was also examined in culture for their characteristic proliferation.

4.5 Overexpression of oncogenes can lead stem cells and/or luminal progenitors to be the putative TICs *in vitro*

In combination with the tractable oncogenic mouse mammary tumor models, we are able to examine the response of adult mammary epithelial populations undergo overexpression of oncogenes *in vivo*. Upon overexpression of Neu, unexpectedly, no significant differences were observed prior to and post induction among the luminal progenitor population. We suspect that those cells are highly dependent on the signaling from their surrounding microenvironment. Our *in vitro* suggests that the contribution of stroma that lacks in the culture system is a major contributor to oncogenic transformation of luminal progenitor cells.

We observed heterogeneity in response to oncogene overexpression among the stem cell population enriched by FACS. Almost 30% of stem cell colonies expanded and gave rise to an aberrant morphological phenotype with an increased proliferation, but the rest showed no apparent differences. This observation and percentage of aberrant colonies onDox were consistent in the case of both Neu and Myc overexpression among the stem cells in culture. This could be due to the heterogeneous expression of MMTV-rtTA promoter. In fact, one of the reports claimed that MMTV-rtTA expression is actually limited to the luminal cells[163]. However, our transgene expression analysis demonstrated that all the FACS enriched populations including myoepithelial and stem cell lineage expressed MMTV confirmed by qRT-PCR and/or immuno-staining. It is striking that even after only 24 hours of doxycycline induction in culture media, there was a potent overexpression of oncogenes.

4.6 Effect of oncogene for lineage commitment, trans-differentiation or de-differentiation

Using those lineage specific markers K8/18, SMA, K14, and Slug, we further analyzed cellular lineage composition of the onDox aberrant colonies. Upon Neu overexpression, we observed an expansion of luminal cells in the inner part of stem cell colonies. The selective luminal lineage expansion may imply that stem cells are TICs of Neu tumors since Neu driven tumor are composed of luminal cells.

In contrast, there was a gradual reduction or loss of the differentiated myoepithelial marker, SMA, in the outer layer of colonies over a period of time (around 3 weeks). We questioned if those myoepithelial cells disappear by apoptosis. Surprisingly, these outer cells were negative for Caspase-3 dependent apoptosis, maintaining their myoepithelial/basal lineage shown by mild Slug expression, and an un-differentiated myoepithelial cellular marker, K14 expression. Recent report showed that Slug is expressed among the proliferating myoepithelial cells (CD49f, CK5 positive)[114] and Slug down-regulation leads to expansion of the luminal-committed lineage[84,158,164]. Indeed the inner part of stem cell colonies onDox for 3 weeks exhibited a strong K8/18 expression and that were Slug negative confirming their true luminal characteristics. The outer layer of cells moderately expressed Slug and K14 but not SMA. This hints us de-differentiation of myoepithelial-committed lineage towards un-committed myoepithelial progenitors by Neu overexpression. However, we observed no significant effect of oncogene Neu in luminal progenitor population in culture.

On the other hand, Myc overexpression gave significant aberrant morphological changes and increases in proliferation upon culture in the majority of luminal progenitors and around 30% of stem cells. We observed classical features like the loss of epithelial polarity and the formation of solid spheres of luminal progenitor cells as described previously using the tractable Myc oncogene overexpression combined with Kras[154]. Loss of normal tissue organization is considered as one of the early diagnostic features of ductal carcinoma *in situ*[165,166] and demonstrates the tumor promoting effect of Myc alone in the tested cell types. In contrast to Neu, overexpression of Myc maintained the myoepithelial lineage as shown by SMA expression and diminished luminal K8/18 positive cells within the stem cell population. Luminal progenitors, however, expanded the cell types of both linages and kept a substantial K8/18 positivity in the luminal fraction, making them optimal candidates as TICs for the Myc oncogene. In more detail, we also observed de-differentiation and trans-differentiation of Luminal progenitors to committed myoepithelial lineage expressing SMA in the outer layer of their colonies. We again confirmed co-expression of SMA and Slug supporting that true myoepithelial lineage commitment as described previously[114]. We have observed that distinct oncogenic mutations can alter the physiological cell-fate decision. All of those data together hints that in the presence of an active oncogene, classical cellular hierarchies are not strictly followed. We would further like to confirm the lineage transition induced by oncogenes. Currently two experiments to further confirm these findings are ongoing:

1) Continuation of transplantation experiment of H2B-GFP positive T-O-NEU or T-O-MYC for tumorigenesis. This would further clarify the transition of cellular lineage and contribution to tumor phenotypes especially in the case of Myc overexpressing luminal progenitor trans-differentiation.
2) Another ongoing experiment *in vivo* is the transplantation of mixed population i.e. T-O-NEU or T-O-MYC cell population together with wild type non-FACS enriched cells. This may reduce the possible artifactual trans-differentiation from the transplantation process and contribution of oncogene overexpressing cells toward tumorigenesis can be clearly defined.

4.7 Different cell types such as stem cells and luminal progenitor cells can be TICs but not the differentiated cells

Myo low population consists nearly half (40-45%) of all mammary epithelial cells. Our *in vitro* culture of different FACS enriched populations indicated that Myo low (stromal cells or myoepithelial differentiated cells) populations are much less responsive to the oncogene overexpression when measured by the aberrance of proliferation and morphology. More rare populations such as stem cells, luminal progenitor cells presented aberrant phenotypes upon overexpression of oncogenes. However, the transgene expression is potent in Myo low population and greater in the case of Myc. From our *in vivo* tumorigenesis assay, we observed tumor formation in different cell populations including stem cells and luminal progenitor cells. In case of both Neu and Myc, Myo low population rarely developed tumors upon over 2000 cells is injected. Our *in vivo* results indicate that different subpopulation within the mammary epithelial hierarchy can serve as TICs except Myo low population (fully-differentiated cells and stromal cells). What we define as differentiated cells in this study are non-stem and non-progenitor cell population, however, it is necessary to further enrich distinct functionally differentiated mammary epithelial cells. In the future, we would like to investigate further why those differentiated cells are less susceptible to overexpression of oncogenes including their proliferation and epigenetic status.

Both tumors arising from T-O-NEU stem cell and luminal progenitor cell transplantations showed histo-pathologically close similarity to that of primary Neu overexpressing tumors. We could not identify the major differences in tumors arising from different cell population assessed by histological phenotypes or cellular compositions. Similarly, Myc overexpression led to development of tumors from the

transplantation of different subpopulations. More frequently, myoepithelial population including the stem cells and Myo hi gave rise to Myc driven tumor formation as well as some luminal progenitors. In both cases, we did not observe any significant differences between the tumors arising from different cell populations, we suspect that tumor phenotypes could be determined by the driving oncogene rather than tumor cell-of-origin in case of Neu and Myc overexpression. In other words, regardless of the cell-of-origin, the genetic mutation could determine the tumor in the specific breast cancer subtype.

4.8 Mesenchymal-like state and epithelial-like state of breast cancer cells

Wicha and colleagues recently demonstrated that the putative breast CSCs exist in both mesenchymal-like state and epithelial-like state that is interchangeable[167]. The mesenchymal-like state is associated with the expression of myoepithelial/basal lineage specific markers (e.g. K14, SMA, CD49f), mesenchymal specific markers such as vimentin, relative quiescence, and highly invasive capacity. In contrast, the epithelial-like state is associated with expression of luminal lineage markers (e.g. K8/18, CD24, EpCAM), E-cadherin (establishment of cell polarity) as well as enhanced proliferation. In this study, both genetic and micro-environmental factors are suggested to be the major contributors of cellular lineage transition.

In addition, their study also implicates that different breast cancer molecular subtypes are characterized by distinct genetic mutations. Their findings are reminiscent to our results that show a remarkable plasticity of the cell types when exposed to oncogenic signals. While it is not too surprising that stem cells are able to preferentially expand specific lineage in culture upon overexpression of two contrasting oncogenes. Additionally, we observe plasticity and indication of cellular re-programming in the luminal progenitor population upon Myc overexpression. In the case of MYC overexpression, these findings are not only shown *in vitro*, but also our preliminary data on the *in vivo tumorigenesis* experiments. In addition, BRCA-1 mutation are also reported to induce the liminal progenitors to trans-differentiate to the myoepithelial lineage giving rise to basal-like tumors[84].

Together our finding suggests that the lineage committed luminal progenitor cells were able to de-differentiate into somewhat closer to the multi-potent state and switch their lineage upon overexpression of Myc. This might explain how the luminal progenitors give rise to mixed basal tumor phenotype. This plasticity is limited to the stem cell and luminal progenitor populations in our hands, since terminally

differentiated cells were unable to expand in the face of an oncogene or trans-differentiate.

4.9 Use of lineage tracing approach to identify the Tumor Initiating Cells

In addition to the cell surface proteins that allow us to FACS isolate different mammary epithelial cells, some genetic markers for mammary epithelial cells were reported including several Wnt targeted genes including Axin2[131], and Protein C Receptor[130]. Axin2 is reported as the genetic marker for bi-potent adult mammary stem cells that reside in myoepithelial population[131]. Interestingly, Axin2 positive cells at different stages of development contributed to the formation of luminal cells (marked at embryonic stage E14) and myoepithelial cells (marked at puberty 5 weeks) at the adult stage. Using Axin2-Cre-ERT2/Rosa26-Brainbow2.1 and Axin2-Cre-ERT2/Rosa26-YFP, we first assessed our marking efficiency for lineage tracing. Further cross of these tracing lines with T-O-MYC and T-O-NEU tumorigenic line, we would like to examine the contribution of specific cellular lineage to the formation of mammary tumor. It is possible that efficiency of marking can be problematic. Our initial tracing experiment showed that there are substantial amount of Axin2-YFP positive cells within the Myo low population after 1 week of tamoxifen induction. This was our initial concern of the model although our previous experiments *in vitro* and *in vivo* indicate that Myo low cells are less likely to be TIC giving rise to tumors upon oncogene induction. It was more encouraging that the percentage of YFP positive cells in Myo low stayed almost the same 3 weeks later, whereas YFP positive cells within the stem and Myo hi gating significantly expanded. Despite of the potential concerns, this approach could compliment our *in vivo* transplantation assay data with a closer physiological situation. In the case of very tight regulation and tightness of the marking system, we would overcome the current problems of 1) transplantation artifact *in vivo* and 2) lack of microenvironment in culture. In addition to the contribution of adult stem cells in Neu and Myc induced tumorigenesis, we would like to assess the role of luminal and myoepithelial cells that can be specifically marked at different mammary gland developmental stages.

4.10 Possible consequences for clinical practice

A number of studies and clinical trials are aiming at eradicating the CSCs through targeting their characteristic features including cell surface markers, signaling cascades, micro-environment, ABC cassette that are selective features for CSCs[168]. A study on chronic myeloid leukemia (CML) by Essers *et al.* proposed an example of

CSC specific therapy[169]. Dormant hematopoietic stem cells are known to serve as the CSCs in CML and resistant to the conventional chemotherapy. However, combination therapy of Interferon-alpha (IFN-α) pre-treatment that induce active proliferation of CSCs and anti-proliferation chemotherapeutic agent 5-fluoro-uracil (5FU) increased the treatment efficiency significantly. When the tumors are dependent upon their CSCs for their growth, the CSC targeted therapy can be an effective approach.

What our study might add to the field is the differentiation status of cells at the pre-tumorigenesis stage may play a role upon acquisition of oncogenic mutations. Although it is preliminary, our *in vitro* and *in vivo* results indicate that the differentiated cells within the mammary epithelial hierarchy did not exhibit any morphological aberrance in culture and are less likely to initiate tumors upon transplantation. Having been multi-potent and/or possessing the proliferative capacity might make the cells to be more susceptible to oncogenic hits. This finding supports the clinical relevance of developing so called “differentiation therapy”. Instead of aiming at the eradication of tumor cells, differentiation therapy force to induce the terminal differentiation of un-differentiated tumor cells so that they become less responsive to oncogenic mutation signaling. Previous studies on acute pro-myelocytic leukemia with retinoic acids (vitamin A analogs) are the examples of differentiation therapy by unblocking the maturation arrest[170,171]. A recent report suggested that the use of Cripto antagonist may induce mammary cellular differentiation and reduces self-renewal of mammary stem cells[172]. The small molecules that could de-induced pluripotency or determine the cellular lineage commitment and differentiation in the field of iPS research might open up the new avenue of differentiation therapy in the breast cancer treatment[173,174].

5. Literature

1 Bray, F., Ren, J. S., Masuyer, E. & Ferlay, J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *International journal of cancer. Journal international du cancer* **132**, 1133-1145, doi:10.1002/ijc.27711 (2013).

2 Ferlay J, S. I., Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. . GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 (2013).

3 Ferlay, J. *et al.* Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer. Journal international du cancer* **136**, E359-386, doi:10.1002/ijc.29210 (2015).

4 Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International journal of cancer. Journal international du cancer* **127**, 2893-2917, doi:10.1002/ijc.25516 (2010).

5 Cook, M. B. *et al.* Tobacco and Alcohol in Relation to Male Breast Cancer: An Analysis of the Male Breast Cancer Pooling Project Consortium. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, doi:10.1158/1055-9965.EPI-14-1009 (2014).

6 Weigelt, B. & Reis-Filho, J. S. Histological and molecular types of breast cancer: is there a unifying taxonomy? *Nature reviews. Clinical oncology* **6**, 718-730, doi:10.1038/nrclinonc.2009.166 (2009).

7 Pfefferle, A. D. *et al.* Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. *Genome biology* **14**, R125, doi:10.1186/gb-2013-14-11-r125 (2013).

8 Prat, A. & Perou, C. M. Deconstructing the molecular portraits of breast cancer. *Molecular oncology* **5**, 5-23, doi:10.1016/j.molonc.2010.11.003 (2011).

9 Burrell, R. A., McGranahan, N., Bartek, J. & Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338-345, doi:10.1038/nature12625 (2013).

10 Magee, J. A., Piskounova, E. & Morrison, S. J. Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer cell* **21**, 283-296, doi:10.1016/j.ccr.2012.03.003 (2012).

11 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).

12 Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

13 Osborne, C., Wilson, P. & Tripathy, D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *The oncologist* **9**, 361-377 (2004).

14 Cuny, M. *et al.* Relating genotype and phenotype in breast cancer: an analysis of the prognostic significance of amplification at eight different genes or loci and of p53 mutations. *Cancer Res* **60**, 1077-1083 (2000).

15 Bolufer, P. *et al.* Estradiol receptors in combination with neu or myc oncogene amplifications might define new subtypes of breast cancer. *Clinica chimica acta; international journal of clinical chemistry* **229**, 107-122 (1994).

16 Gaffey, M. J., Frierson, H. F., Jr. & Williams, M. E. Chromosome 11q13, c-erbB-2, and c-myc amplification in invasive breast carcinoma: clinicopathologic correlations. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **6**, 654-659 (1993).

17 Deming, S. L., Nass, S. J., Dickson, R. B. & Trock, B. J. C-myc amplification in breast cancer: a meta-analysis of its occurrence and prognostic relevance. *Br J Cancer* **83**, 1688-1695, doi:10.1054/bjoc.2000.1522 (2000).

18 Liao, D. J. & Dickson, R. B. c-Myc in breast cancer. *Endocrine-related cancer* **7**, 143-164 (2000).
19 Arnold, A. & Papanikolaou, A. Cyclin D1 in breast cancer pathogenesis. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **23**, 4215-4224, doi:10.1200/JCO.2005.05.064 (2005).
20 Santarius, T., Shipley, J., Brewer, D., Stratton, M. R. & Cooper, C. S. A census of amplified and overexpressed human cancer genes. *Nature reviews. Cancer* **10**, 59-64, doi:10.1038/nrc2771 (2010).
21 Eckert, L. B. *et al.* Involvement of Ras activation in human breast cancer cell signaling, invasion, and anoikis. *Cancer Res* **64**, 4585-4592, doi:10.1158/0008-5472.CAN-04-0396 (2004).
22 Markman, B., Atzori, F., Perez-Garcia, J., Tabernero, J. & Baselga, J. Status of PI3K inhibition and biomarker development in cancer therapeutics. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **21**, 683-691, doi:10.1093/annonc/mdp347 (2010).
23 Stephens, P. J. *et al.* The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**, 400-404, doi:10.1038/nature11017 (2012).
24 Fackenthal, J. D. & Olopade, O. I. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nature reviews. Cancer* **7**, 937-948, doi:10.1038/nrc2054 (2007).
25 Nevanlinna, H. & Bartek, J. The CHEK2 gene and inherited breast cancer susceptibility. *Oncogene* **25**, 5912-5919, doi:10.1038/sj.onc.1209877 (2006).
26 Pharoah, P. D., Day, N. E. & Caldas, C. Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *Br J Cancer* **80**, 1968-1973, doi:10.1038/sj.bjc.6690628 (1999).
27 Ahmed, M. & Rahman, N. ATM and breast cancer susceptibility. *Oncogene* **25**, 5906-5911, doi:10.1038/sj.onc.1209873 (2006).
28 Zhang, H. Y., Liang, F., Jia, Z. L., Song, S. T. & Jiang, Z. F. PTEN mutation, methylation and expression in breast cancer patients. *Oncol Lett* **6**, 161-168, doi:Doi 10.3892/Ol.2013.1331 (2013).
29 Herschkowitz, J. I., He, X., Fan, C. & Perou, C. M. The functional loss of the retinoblastoma tumour suppressor is a common event in basal-like and luminal B breast carcinomas. *Breast cancer research : BCR* **10**, R75, doi:10.1186/bcr2142 (2008).
30 Nair, R. *et al.* c-Myc and Her2 cooperate to drive a stem-like phenotype with poor prognosis in breast cancer. 1-11, doi:10.1038/onc.2013.368 (2013).
31 Berger, M. S. *et al.* Correlation of c-erbB-2 gene amplification and protein expression in human breast carcinoma with nodal status and nuclear grading. *Cancer Res* **48**, 1238-1243 (1988).
32 Slamon, D. J. *et al.* Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235**, 177-182 (1987).
33 Coussens, L. *et al.* Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. *Science* **230**, 1132-1139 (1985).
34 Glynn, R. W., Miller, N., Mahon, S. & Kerin, M. J. Expression levels of HER2/neu and those of collocated genes at 17q12-21, in breast cancer. *Oncology reports* **28**, 365-369, doi:10.3892/or.2012.1780 (2012).
35 Rubin, I. & Yarden, Y. The basic biology of HER2. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **12 Suppl 1**, S3-8 (2001).
36 Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70, doi:10.1038/nature11412 (2012).
37 Xu, J., Chen, Y. & Olopade, O. I. MYC and Breast Cancer. *Genes & cancer* **1**, 629-640, doi:10.1177/1947601910378691 (2010).

38 Dang, C. V. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol Cell Biol* **19**, 1-11 (1999).
39 Chandriani, S. *et al.* A core MYC gene expression signature is prominent in basal-like breast cancer but only partially overlaps the core serum response. *PLoS One* **4**, e6693, doi:10.1371/journal.pone.0006693 (2009).
40 Grushko, T. A. *et al.* MYC is amplified in BRCA1-associated breast cancers. *Clinical cancer research : an official journal of the American Association for Cancer Research* **10**, 499-507 (2004).
41 Dowsett, M. *et al.* Meta-analysis of breast cancer outcomes in adjuvant trials of aromatase inhibitors versus tamoxifen. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 509-518, doi:10.1200/JCO.2009.23.1274 (2010).
42 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
43 Weinstein, I. B. & Joe, A. Oncogene addiction. *Cancer Res* **68**, 3077-3080; discussion 3080, doi:10.1158/0008-5472.CAN-07-3293 (2008).
44 Nahta, R. & Esteva, F. J. HER-2-targeted therapy: lessons learned and future directions. *Clinical cancer research : an official journal of the American Association for Cancer Research* **9**, 5078-5084 (2003).
45 Harari, D. & Yarden, Y. Molecular mechanisms underlying ErbB2/HER2 action in breast cancer. *Oncogene* **19**, 6102-6114, doi:10.1038/sj.onc.1203973 (2000).
46 Sliwkowski, M. X. *et al.* Nonclinical studies addressing the mechanism of action of trastuzumab (Herceptin). *Seminars in oncology* **26**, 60-70 (1999).
47 Cho, H. S. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* **421**, 756-760, doi:10.1038/nature01392 (2003).
48 Kumler, I., Tuxen, M. K. & Nielsen, D. L. A systematic review of dual targeting in HER2-positive breast cancer. *Cancer treatment reviews* **40**, 259-270, doi:10.1016/j.ctrv.2013.09.002 (2014).
49 Baselga, J. *et al.* Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. *The New England journal of medicine* **366**, 109-119, doi:10.1056/NEJMoa1113216 (2012).
50 Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* **3**, 730-737 (1997).
51 Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645-648, doi:10.1038/367645a0 (1994).
52 Politi, K. & Pao, W. How genetically engineered mouse tumor models provide insights into human cancers. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **29**, 2273-2281, doi:10.1200/JCO.2010.30.8304 (2011).
53 Stewart, T. A., Pattengale, P. K. & Leder, P. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell* **38**, 627-637 (1984).
54 Donehower, L. A. *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215-221, doi:10.1038/356215a0 (1992).
55 Jacks, T. *et al.* Tumor spectrum analysis in p53-mutant mice. *Current biology : CB* **4**, 1-7 (1994).
56 Muller, W. J., Sinn, E., Pattengale, P. K., Wallace, R. & Leder, P. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. *Cell* **54**, 105-115 (1988).

57 Guy, C. T. *et al.* Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 10578-10582 (1992).
58 Schoenenberger, C. A. *et al.* Targeted c-myc gene expression in mammary glands of transgenic mice induces mammary tumours with constitutive milk protein gene transcription. *EMBO J* **7**, 169-175 (1988).
59 Wagner, K. U. *et al.* Cre-mediated gene deletion in the mammary gland. *Nucleic acids research* **25**, 4323-4330 (1997).
60 Selbert, S. *et al.* Efficient BLG-Cre mediated gene deletion in the mammary gland. *Transgenic Res* **7**, 387-396 (1998).
61 Chin, L. *et al.* Essential role for oncogenic Ras in tumour maintenance. *Nature* **400**, 468-472, doi:10.1038/22788 (1999).
62 Felsher, D. W. & Bishop, J. M. Reversible tumorigenesis by MYC in hematopoietic lineages. *Mol Cell* **4**, 199-207 (1999).
63 Pelengaris, S., Littlewood, T., Khan, M., Elia, G. & Evan, G. Reversible activation of c-Myc in skin: induction of a complex neoplastic phenotype by a single oncogenic lesion. *Mol Cell* **3**, 565-577 (1999).
64 Zhu, Z., Zheng, T., Lee, C. G., Homer, R. J. & Elias, J. A. Tetracycline-controlled transcriptional regulation systems: advances and application in transgenic animal modeling. *Seminars in cell & developmental biology* **13**, 121-128 (2002).
65 Jones, J. M. & Gellert, M. The taming of a transposon: V(D)J recombination and the immune system. *Immunological reviews* **200**, 233-248, doi:10.1111/j.0105-2896.2004.00168.x (2004).
66 Taneja, P. *et al.* Transgenic and knockout mice models to reveal the functions of tumor suppressor genes. *Clinical Medicine Insights. Oncology* **5**, 235-257, doi:10.4137/CMO.S7516 (2011).
67 Feil, R., Wagner, J., Metzger, D. & Chambon, P. Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem Biophys Res Commun* **237**, 752-757, doi:10.1006/bbrc.1997.7124 (1997).
68 Kretzschmar, K. & Watt, F. M. Lineage tracing. *Cell* **148**, 33-45, doi:10.1016/j.cell.2012.01.002 (2012).
69 Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003-1007, doi:10.1038/nature06196 (2007).
70 Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21**, 70-71, doi:10.1038/5007 (1999).
71 Shibata, H. *et al.* Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. *Science* **278**, 120-123 (1997).
72 Barker, N. *et al.* Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608-612, doi:10.1038/nature07602 (2009).
73 Giel-Moloney, M., Krause, D. S., Chen, G., Van Etten, R. A. & Leiter, A. B. Ubiquitous and uniform in vivo fluorescence in ROSA26-EGFP BAC transgenic mice. *Genesis* **45**, 83-89, doi:10.1002/dvg.20269 (2007).
74 Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC developmental biology* **1**, 4 (2001).
75 Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56-62, doi:10.1038/nature06293 (2007).
76 Snippert, H. J. *et al.* Intestinal Crypt Homeostasis Results from Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells. *Cell* **143**, 134-144, doi:10.1016/j.cell.2010.09.016 (2010).
77 Visvader, J. E. Cells of origin in cancer. *Nature* **469**, 314-322, doi:10.1038/nature09781 (2011).

78 Fu, N., Lindeman, G. J. & Visvader, J. E. The mammary stem cell hierarchy. *Current topics in developmental biology* **107**, 133-160, doi:10.1016/B978-0-12-416022-4.00005-6 (2014).
79 Rapp, U. R., Ceteci, F. & Schreck, R. Oncogene-induced plasticity and cancer stem cells. *Cell cycle* **7**, 45-51 (2008).
80 Abeliovich, D. *et al.* The founder mutations 185delAG and 5382insC in BRCA1 and 6174delT in BRCA2 appear in 60% of ovarian cancer and 30% of early-onset breast cancer patients among Ashkenazi women. *American journal of human genetics* **60**, 505-514 (1997).
81 Nicoletto, M. O. *et al.* BRCA-1 and BRCA-2 mutations as prognostic factors in clinical practice and genetic counselling. *Cancer treatment reviews* **27**, 295-304, doi:10.1053/ctrv.2001.0233 (2001).
82 Lim, E. *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15**, 907-913, doi:10.1038/nm.2000 (2009).
83 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, doi:10.1016/j.stem.2010.07.010 (2010).
84 Proia, T. A. *et al.* Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. *Cell Stem Cell* **8**, 149-163, doi:10.1016/j.stem.2010.12.007 (2011).
85 Wagner, K. U., Booth, B. W., Boulanger, C. A. & Smith, G. H. Multipotent PI-MECs are the true targets of MMTV-neu tumorigenesis. *Oncogene* **32**, 1338, doi:10.1038/onc.2012.452 (2013).
86 Lo, P. K. & Chen, H. Cancer stem cells and cells of origin in MMTV-Her2/neu-induced mammary tumorigenesis. *Oncogene* **32**, 1338-1340, doi:10.1038/onc.2012.456 (2013).
87 Henry, M. D., Triplett, A. A., Oh, K. B., Smith, G. H. & Wagner, K. U. Parity-induced mammary epithelial cells facilitate tumorigenesis in MMTV-neu transgenic mice. *Oncogene* **23**, 6980-6985, doi:10.1038/sj.onc.1207827 (2004).
88 Matulka, L. A., Triplett, A. A. & Wagner, K. U. Parity-induced mammary epithelial cells are multipotent and express cell surface markers associated with stem cells. *Dev Biol* **303**, 29-44, doi:10.1016/j.ydbio.2006.12.017 (2007).
89 Jeselsohn, R. *et al.* Cyclin D1 Kinase Activity Is Required for the Self-Renewal of Mammary Stem and Progenitor Cells that Are Targets of MMTV-ErbB2 Tumorigenesis. *Cancer cell* **17**, 65-76, doi:10.1016/j.ccr.2009.11.024 (2010).
90 Liu, J. C., Deng, T., Lehal, R. S., Kim, J. & Zacksenhaus, E. Identification of Tumorsphere- and Tumor-Initiating Cells in HER2/Neu-Induced Mammary Tumors. *Cancer Research* **67**, 8671-8681, doi:10.1158/0008-5472.CAN-07-1486 (2007).
91 Lo, P. K. *et al.* CD49f and CD61 identify Her2/neu-induced mammary tumor-initiating cells that are potentially derived from luminal progenitors and maintained by the integrin-TGFbeta signaling. *Oncogene* **31**, 2614-2626, doi:10.1038/onc.2011.439 (2012).
92 Andrechek, E. R. *et al.* Genetic heterogeneity of Myc-induced mammary tumors reflecting diverse phenotypes including metastatic potential. **106**, 16387-16392, doi:10.1073/pnas.0901250106 (2009).
93 Leung, J. Y., Andrechek, E. R., Cardiff, R. D. & Nevins, J. R. Heterogeneity in MYC-induced mammary tumors contributes to escape from oncogene dependence. **31**, 2545-2554, doi:10.1038/onc.2011.433 (2011).
94 Mikaelian, I. *et al.* Expression of terminal differentiation proteins defines stages of mouse mammary gland development. *Veterinary pathology* **43**, 36-49, doi:10.1354/vp.43-1-36 (2006).

95 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, doi:10.1016/j.stem.2011.12.018 (2012).
96 Makarem, M. *et al.* Stem cells and the developing mammary gland. *J Mammary Gland Biol Neoplasia* **18**, 209-219, doi:10.1007/s10911-013-9284-6 (2013).
97 Zeps, N., Bentel, J. M., Papadimitriou, J. M., D'Antuono, M. F. & Dawkins, H. J. Estrogen receptor-negative epithelial cells in mouse mammary gland development and growth. *Differentiation; research in biological diversity* **62**, 221-226, doi:10.1046/j.1432-0436.1998.6250221.x (1998).
98 Joshi, P. A. *et al.* Progesterone induces adult mammary stem cell expansion. *Nature* **465**, 803-807, doi:10.1038/nature09091 (2010).
99 Schweizer, J. *et al.* New consensus nomenclature for mammalian keratins. *J Cell Biol* **174**, 169-174, doi:10.1083/jcb.200603161 (2006).
100 Iyer, S. V. *et al.* Understanding the role of keratins 8 and 18 in neoplastic potential of breast cancer derived cell lines. *PLoS One* **8**, e53532, doi:10.1371/journal.pone.0053532 (2013).
101 Paramio, J. M. & Jorcano, J. L. Beyond structure: do intermediate filaments modulate cell signalling? *BioEssays : news and reviews in molecular, cellular and developmental biology* **24**, 836-844, doi:10.1002/bies.10140 (2002).
102 Coulombe, P. A. & Omary, M. B. 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. *Curr Opin Cell Biol* **14**, 110-122 (2002).
103 Owens, D. W. & Lane, E. B. The quest for the function of simple epithelial keratins. *BioEssays : news and reviews in molecular, cellular and developmental biology* **25**, 748-758, doi:10.1002/bies.10316 (2003).
104 Oshima, R. G. Apoptosis and keratin intermediate filaments. *Cell Death Differ* **9**, 486-492, doi:10.1038/sj/cdd/4400988 (2002).
105 Baribault, H., Wilson-Heiner, M., Muller, W., Penner, J. & Bakhiet, N. Functional analysis of mouse keratin 8 in polyoma middle T-induced mammary gland tumours. *Transgenic Res* **6**, 359-367 (1997).
106 Sopel, M. The myoepithelial cell: its role in normal mammary glands and breast cancer. *Folia morphologica* **69**, 1-14 (2010).
107 Gudjonsson, T. *et al.* Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition. *J Cell Sci* **115**, 39-50 (2002).
108 Wetzels, R. H. *et al.* Basal cell-specific and hyperproliferation-related keratins in human breast cancer. *Am J Pathol* **138**, 751-763 (1991).
109 Han, A. C., Soler, A. P., Knudsen, K. A. & Salazar, H. Distinct cadherin profiles in special variant carcinomas and other tumors of the breast. *Human pathology* **30**, 1035-1039 (1999).
110 Daniel, C. W., Strickland, P. & Friedmann, Y. Expression and functional role of E- and P-cadherins in mouse mammary ductal morphogenesis and growth. *Dev Biol* **169**, 511-519, doi:10.1006/dbio.1995.1165 (1995).
111 Carroll, D. K. *et al.* p63 regulates an adhesion programme and cell survival in epithelial cells. *Nat Cell Biol* **8**, 551-561, doi:10.1038/ncb1420 (2006).
112 Barbareschi, M. *et al.* p63, a p53 homologue, is a selective nuclear marker of myoepithelial cells of the human breast. *The American journal of surgical pathology* **25**, 1054-1060 (2001).
113 Chakrabarti, R. *et al.* Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. *Stem Cells* **30**, 1496-1508, doi:10.1002/stem.1112 (2012).
114 Nassour, M. *et al.* Slug controls stem/progenitor cell growth dynamics during mammary gland morphogenesis. *PLoS One* **7**, e53498, doi:10.1371/journal.pone.0053498 (2012).

115 Till, J. E. & Mc, C. E. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation research* **14**, 213-222 (1961).
116 Siminovitch, L., McCulloch, E. A. & Till, J. E. The Distribution of Colony-Forming Cells among Spleen Colonies. *J Cell Physiol* **62**, 327-336 (1963).
117 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).
118 Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84-88, doi:10.1038/nature04372 (2006).
119 Visvader, J. E. & Stingl, J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes & development* **28**, 1143-1158, doi:10.1101/gad.242511.114 (2014).
120 Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature*, 1-39, doi:10.1038/nature10573 (2011).
121 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, doi:10.1038/nature12948 (2014).
122 Oakes, S. R., Gallego-Ortega, D. & Ormandy, C. J. The mammary cellular hierarchy and breast cancer. *Cellular and molecular life sciences : CMLS* **71**, 4301-4324, doi:10.1007/s00018-014-1674-4 (2014).
123 Visvader, J. E. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. *Genes & Development* **23**, 2563-2577, doi:10.1101/gad.1849509 (2009).
124 Kalirai, H. & Clarke, R. B. Human breast epithelial stem cells and their regulation. *J Pathol* **208**, 7-16, doi:10.1002/path.1881 (2006).
125 Cicalese, A. *et al.* The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell* **138**, 1083-1095, doi:10.1016/j.cell.2009.06.048 (2009).
126 Pece, S. *et al.* Biological and Molecular Heterogeneity of Breast Cancers Correlates with Their Cancer Stem Cell Content. *Cell* **140**, 62-73, doi:10.1016/j.cell.2009.12.007 (2010).
127 Ginestier, C. *et al.* ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **1**, 555-567, doi:10.1016/j.stem.2007.08.014 (2007).
128 Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature Cell Biology*, doi:10.1038/nature04496 (2006).
129 Prater, M. D. *et al.* Mammary stem cells have myoepithelial cell properties. *Nat Cell Biol* **16**, 942-950, doi:10.1038/ncb3025 (2014).
130 Wang, D. *et al.* Identification of multipotent mammary stem cells by protein C receptor expression. *Nature*, doi:10.1038/nature13851 (2014).
131 van Amerongen, R., Bowman, A. N. & Nusse, R. Developmental stage and time dictate the fate of Wnt/beta-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* **11**, 387-400, doi:10.1016/j.stem.2012.05.023 (2012).
132 dos Santos, C. O. *et al.* Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 7123-7130, doi:10.1073/pnas.1303919110 (2013).
133 Asselin-Labat, M. L. *et al.* Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat Cell Biol* **9**, 201-209, doi:10.1038/ncb1530 (2007).
134 Shehata, M. *et al.* Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. *Breast cancer research : BCR* **14**, R134, doi:10.1186/bcr3334 (2012).

135 Asselin-Labat, M. L. *et al.* Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. *Mol Cell Biol* **31**, 4609-4622, doi:10.1128/MCB.05766-11 (2011).
136 Regan, J. L. *et al.* c-Kit is required for growth and survival of the cells of origin of Brca1-mutation-associated breast cancer. *Oncogene* **31**, 869-883, doi:10.1038/onc.2011.289 (2012).
137 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, doi:10.1083/jcb.200604065 (2007).
138 Hu, Y. & Smyth, G. K. ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *Journal of Immunological Methods* **347**, 70-78, doi:http://dx.doi.org/10.1016/j.jim.2009.06.008 (2009).
139 Daniel, C. W., De Ome, K. B., Young, J. T., Blair, P. B. & Faulkin, L. J., Jr. The in vivo life span of normal and preneoplastic mouse mammary glands: a serial transplantation study. *Proceedings of the National Academy of Sciences of the United States of America* **61**, 53-60 (1968).
140 Schmeichel, K. L. & Bissell, M. J. Modeling tissue-specific signaling and organ function in three dimensions. *J Cell Sci* **116**, 2377-2388, doi:10.1242/jcs.00503 (2003).
141 Jechlinger, M., Podsypanina, K. & Varmus, H. Regulation of transgenes in three-dimensional cultures of primary mouse mammary cells demonstrates oncogene dependence and identifies cells that survive deinduction. *Genes & development* **23**, 1677-1688, doi:10.1101/gad.1801809 (2009).
142 Badders, N. M. *et al.* The Wnt receptor, Lrp5, is expressed by mouse mammary stem cells and is required to maintain the basal lineage. *PLoS One* **4**, e6594, doi:10.1371/journal.pone.0006594 (2009).
143 Bai, L. & Rohrschneider, L. R. s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes & development* **24**, 1882-1892, doi:10.1101/gad.1932810 (2010).
144 Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189-193, doi:10.1038/nature10573 (2011).
145 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, doi:10.1002/path.4096 (2012).
146 Plaks, V. *et al.* Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. *Cell reports* **3**, 70-78, doi:10.1016/j.celrep.2012.12.017 (2013).
147 D'Cruz, C. M. *et al.* c-MYC induces mammary tumorigenesis by means of a preferred pathway involving spontaneous Kras2 mutations. *Nat Med* **7**, 235-239, doi:10.1038/84691 (2001).
148 Moody, S. E. *et al.* Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis. *Cancer cell* **2**, 451-461 (2002).
149 Hadjantonakis, A. K. & Papaioannou, V. E. Dynamic in vivo imaging and cell tracking using a histone fluorescent protein fusion in mice. *BMC biotechnology* **4**, 33, doi:10.1186/1472-6750-4-33 (2004).
150 Snippert, H. J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134-144, doi:10.1016/j.cell.2010.09.016 (2010).
151 Mombaerts, P. *et al.* RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* **68**, 869-877 (1992).

152 Donehower, L. A., Huang, A. L. & Hager, G. L. Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. *Journal of virology* **37**, 226-238 (1981).
153 Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. In situ identification of bipotent stem cells in the mammary gland. 1-19, doi:10.1038/nature12948 (2014).
154 Jechlinger, M., Podsypanina, K. & Varmus, H. Regulation of transgenes in three-dimensional cultures of primary mouse mammary cells demonstrates oncogene dependence and identifies cells that survive deinduction. *Genes & Development* **23**, 1677-1688, doi:10.1101/gad.1801809 (2009).
155 McLean, A. C., Valenzuela, N., Fai, S. & Bennett, S. A. Performing vaginal lavage, crystal violet staining, and vaginal cytological evaluation for mouse estrous cycle staging identification. *Journal of visualized experiments : JoVE*, e4389, doi:10.3791/4389 (2012).
156 D'cruz. c-MYC induces mammary tumorigenesis by means of a preferred pathway involving spontaneous Kras2 mutations. 1-5 (2001).
157 Siegel, P. M., Ryan, E. D., Cardiff, R. D. & Muller, W. J. Elevated expression of activated forms of Neu/ErbB-2 and ErbB-3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer. *EMBO J* **18**, 2149-2164, doi:10.1093/emboj/18.8.2149 (1999).
158 Phillips, S. *et al.* Cell-State Transitions Regulated by SLUG Are Critical for Tissue Regeneration and Tumor Initiation. *Stem cell reports* **2**, 633-647, doi:10.1016/j.stemcr.2014.03.008 (2014).
159 Proia, T. A. *et al.* Genetic Predisposition Directs Breast Cancer Phenotype by Dictating Progenitor Cell Fate. *Stem Cell* **8**, 149-163, doi:10.1016/j.stem.2010.12.007 (2011).
160 Lo, P.-K. *et al.* CD49f and CD61 identify Her2/neu-induced mammary tumor-initiating cells that are potentially derived from luminal progenitors and maintained by the integrin–TGFβ signaling. *Oncogene* **31**, 2614-2626, doi:10.1038/onc.2011.439 (2011).
161 Guo, W. *et al.* Slug and Sox9 Cooperatively Determine the Mammary Stem Cell State. *Cell* **148**, 1015-1028, doi:10.1016/j.cell.2012.02.008 (2012).
162 Clarke, M. F. *et al.* Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. *Cancer Res* **66**, 9339-9344, doi:10.1158/0008-5472.CAN-06-3126 (2006).
163 Plichta, K. A., Mathers, J. L., Gestl, S. A., Glick, A. B. & Gunther, E. J. Basal but not luminal mammary epithelial cells require PI3K/mTOR signaling for Ras-driven overgrowth. *Cancer Res* **72**, 5856-5866, doi:10.1158/0008-5472.CAN-12-1635 (2012).
164 Tripathi, M. K., Misra, S. & Chaudhuri, G. Negative regulation of the expressions of cytokeratins 8 and 19 by SLUG repressor protein in human breast cells. *Biochem Biophys Res Commun* **329**, 508-515, doi:10.1016/j.bbrc.2005.02.006 (2005).
165 Porter, D. *et al.* Molecular markers in ductal carcinoma in situ of the breast. *Molecular cancer research : MCR* **1**, 362-375 (2003).
166 Burstein, H. J., Polyak, K., Wong, J. S., Lester, S. C. & Kaelin, C. M. Ductal carcinoma in situ of the breast. *The New England journal of medicine* **350**, 1430-1441, doi:10.1056/NEJMra031301 (2004).
167 Liu, S. *et al.* Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem cell reports* **2**, 78-91, doi:10.1016/j.stemcr.2013.11.009 (2014).
168 Chen, K., Huang, Y. H. & Chen, J. L. Understanding and targeting cancer stem cells: therapeutic implications and challenges. *Acta pharmacologica Sinica* **34**, 732-740, doi:10.1038/aps.2013.27 (2013).

169 Essers, M. A. *et al.* IFNalpha activates dormant haematopoietic stem cells in vivo. *Nature* **458**, 904-908, doi:10.1038/nature07815 (2009).
170 Sell, S. Cancer stem cells and differentiation therapy. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **27**, 59-70, doi:10.1159/000092323 (2006).
171 Sell, S. Cellular origin of cancer: dedifferentiation or stem cell maturation arrest? *Environmental health perspectives* **101 Suppl 5**, 15-26 (1993).
172 Spike, B. T. *et al.* CRIPTO/GRP78 signaling maintains fetal and adult mammary stem cells ex vivo. *Stem cell reports* **2**, 427-439, doi:10.1016/j.stemcr.2014.02.010 (2014).
173 Shan, J. *et al.* Identification of small molecules for human hepatocyte expansion and iPS differentiation. *Nature chemical biology* **9**, 514-520, doi:10.1038/nchembio.1270 (2013).
174 Kanke, K. *et al.* Stepwise differentiation of pluripotent stem cells into osteoblasts using four small molecules under serum-free and feeder-free conditions. *Stem cell reports* **2**, 751-760, doi:10.1016/j.stemcr.2014.04.016 (2014).