# ANALYSIS OF MURINE MAMMARY EPITHELIAL CELLULAR HIERARCHY

**by**

**Neethan A. Lobo**

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in The University of Michigan
2009

Doctoral Committee:

Professor Charles Burant, Co-Chair
Professor Michael F. Clarke, Co-Chair, Stanford University
Professor Gregory R. Dressler
Professor J. Douglas Engel
Professor Eric R. Fearon

To my parents Lancy and Diana Lobo and my sister Maria

In loving memory of Laurie Luzynski

## ACKNOWLEDGEMENTS

First, I'd like to thank my mentor, Michael Clarke. He and his family have shown me great kindness throughout the years. I'd also like to thank everyone who made this personal endeavor worth it. For my family and friends, your patience and understanding will never be forgotten. Mom and Dad, you have been extraordinary with all of your support and generosity. I could not have achieved this milestone without you. I want to thank the folks at both University of Michigan and Stanford University, since I basically split my doctorate between the two. To 502 Elm Street, thank you for making graduate school lots of fun.

This work was partially funded by the Department of Defense Breast Cancer Research Program Pre-Doctoral Traineeship Award, grant number DAMD 17-03-1-0340.

## PREFACE

This thesis summarizes my work as of my defense date of April 16, 2009. Dalong Qian aided me in performing many of the *in vivo* transplants covered in chapter 1 for investigating the enrichment of mammary stem cells with the Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ phenotype. Maider Zabala and I isolated the cells used for all microarray analysis described in chapter 3. In addition, Maider Zabala performed the immunocytochemistry for keratin proteins in the mature mammary ducts described in chapter 3. Maximillian Diehn and I performed the DFA-DA staining, flow cytometry and in vivo transplants for the ROS studies in chapter 3.

In my introduction, I cover the concepts of stem cell biology and highlight the importance of studying these powerful cells. There is also an in-depth review of mouse mammary biology since my studies deal exclusively with this organ system. At the end of the introduction, I focus on the elucidation of the mammary hierarchy using cell separation technologies combined with transplantation. In chapter 1, a marker study of Thy-1 shows it is useful in isolating mammary stem cells. By fractionating the existing phenotypic stem cell population based on Thy-1 expression, I propose a new mammary population of multipotent progenitor cells. In my hands, these cells were different than previously described progenitors since they had some ability to make mammary epithelium in vivo, but had severely diminished self-renewal ability.

In chapter 2, I studied how gene expression changes in the mammary hierarchy. Using freely available bioinformatic software, I present a detailed analysis of keratin expression in these populations as a method of comparing progenitor populations to differentiated cells. This work also led to the discovery of a mouse ESA antibody that may be used as a tool to distinguish parenchyma from stroma. I used differentiated populations as filters to screen for genes that were only expressed in the mammary stem cells. When this list was functionally annotated, it revealed new insights about how stem cells interact with their environment.

My bioinformatic analysis allowed prompted me to pursue investigation of signaling pathways known to be expressed in normal cells and often corrupted in cancer. To that end, I studied how ROS levels affect stem cell growth. I present evidence that suggests stem cells with low intracellular ROS have a growth advantage *in vivo*. In keeping with the idea that isolated mammary populations have unique transcriptional profiles, I decided to assess how TGFβ ligand affects individual mammary populations *in vitro*. Although I didn't observe any real effects on mammary stem or multipotent progenitor cells, I did see that TGFβ induced a motility program in early progenitor cells. These results showed that cytokines can have very different effects on cells in the mammary system. Future work on the mammary system should take these differential effects into account.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## ABSTRACT

The murine mammary system is a complex milieu of epithelial cell types that function together to support lactogenesis. The breast tissue is spatially and temporally regulated, with most growth occurring postnatally. The mature mammary gland displays two characteristic architectural features: ducts and terminal end buds. There are different cell types found in these two features that perform independent functions in the mature organ. In general, the mature gland is a bilayer system with a single layer of fibroblast-like myoepithelial cells encompassing an inner luminal epithelial layer. Upon pregnancy, the induction of alveolar secretory cells is initiated, which eventually leads to milk production. Milk is then channeled through the entire system via contraction of the myoepithelium, and eventually dispersed through the nipple. The mouse mammary gland is capable of multiple stages of development: growth, lactation and involution. These stages are regulated by cytokines and hormones depending on the developmental stage of the mouse and the physiological context (i.e. estrus). Numerous studies have shown the mammary epithelium is capable of regeneration upon transplantation. Recent reports have isolated fractions of cells enriched for duct forming ability and self-renewal. In addition, a distinct progenitor population was also identified that lacked the ability to form ducts in vivo but retained bipotentiality *in vitro*. This thesis builds on previous marker studies of mammary epithelial cells, using existing phenotypes to further enrich

the mammary stem cell population. My data shows Thy-1, or CD90, is differentially expressed in mammary cells. In addition, Thy-$1^{+}$CD$24^{med}$CD49$f^{hi}$ cells are enriched for *in vivo* engraftment of ductal epithelium. These cells are approximately 10-fold enriched for duct-forming ability.. Importantly, the Thy-$1^{+}$CD$24^{med}$CD$49^{hi}$ phenotype enriched for cells that are capable of self-renewal and retain the full differentiation potential of the parent stem cells.. This data provided a new phenotype for the stem cells (MaSC). The Thy-$1^{-}$CD$24^{med}$CD$49^{hi}$ cells are not only diminished for engraftment but have significantly decreased self-renewal, suggesting they are a potential multipotent progenitor population. I have named them mammary multipotent progenitors, or MMPPs. A novel protein staining methodology revealed the Thy-$1^{+}$CD$24^{med}$CD$49^{hi}$ and Thy-$1^{-}$ CD$24^{med}$CD$49^{hi}$ cells have keratin profile similar to myoepithelial cells, whereas the further differentiated progenitors initiated a luminal epithelial transcriptional program. In addition, my work shows the first description of murine Epithelial Specific Antigen (or EpCAM) using anantibody produced by the Developmental Studies Hybridoma Bank (NIH, Iowa) in the mammary gland. This antibody successfully discriminates luminal cells from myoepithelial cells, and the staining patterns mirrors ESA in the human mammary system. To further investigate the transcriptional regulation of the mammary populations, I performed microarray analysis of the Thy-$1^{+}$CD$24^{med}$CD$49^{hi}$ (MaSC), Thy-$1^{-}$CD$24^{med}$CD$49^{hi}$(MMPP), MaCFC progenitors, differentiated myoepithelial (MYO) and differentiated luminal epithelial (EPI) cells. This work is the first description of transcriptome analysis of differentiated luminal epithelial cells directly isolated from

fresh mouse tissue. The results show Thy-1$^{+}$CD24$^{med}$CD49$^{hi}$ cells express gene programs consistent with their predicted basal location *in vivo*. In addition, I identified a set of genes that may be used to discriminate stem cells from myoepithelial cells. Discovering gene sets that were specific to hierarchical mammary populations prompted the hypothesis that DNA damage responses and cytokine signaling may also affect those populations in different ways. I went on to investigated how ROS and TGFβ affect the cellular hierarchy. My studies show that stem cells with low intracellular ROS have a growth advantage *in vivo*. In addition, I found that exogenous TGFβ ligand had little effect on MaSC and MMPP colony formation *in vitro*, but did induce a motility phenotype in early progenitor MaCFC cells. In conclusion, my thesis work led to a number of interesting and novel insights into mammary biology. By using the tools described in this thesis, it is now easier to identify which types of cells within the breast are susceptible for transformation. Knowing the cell(s) of origin for breast cancers will be the key to future therapeutic strategies.

# CHAPTER 1

## INTRODUCTION

The discovery that tissues in the human body are not only heterogeneous but composed of a cellular hierarchy has elucidated many new insights into both how we normally function as well as how diseases occur within our bodies. Therefore, stem cell biology has provided a new platform to address many developmental questions. Scientists around the world have been using stem cells as model systems to examine how tissues respond during normal physiology and what happens when normal functions are disturbed. These studies have led to new hypotheses of how cancers are initiated. Taken together, stem cell biology is now cemented as an important field of study.

The body plan of higher organisms follows ordered developmental stages. Typical development follows a pre-determined plan that turns a fertilized egg into a complex, multicellular organism. The fertilized egg develops into a totipotent ball of cells that proceeds through gastrulation which definesthe three germ layers: endoderm, ectoderm and mesoderm[1]. These three primitive cell types develop into all the tissues in the adult body. Cells in the body may be broadly categorized as either germ or somatic cells. Germ cells have the unique ability to undergo both mitosis and meiosis. Although germ cells are long lived, even thought to be immortal with respect to the lifetime of the human from which they come, they are quite different than somatic stem cells since their

purpose is to transmit genetic information rather than maintain the health of the organism. Somatic tissues such as the blood, brain, breast, liver, fat, skin and gut contain a small population of tissue specific stem cells responsible for both development and maintenance of those tissues for the human lifetime. There exists a cellular hierarchy in these tissues, but the regulatory systems that govern proliferation and differentiation are poorly understood. Recent advances in cell separation technologies have enabled researchers to investigate signaling pathways to understand how tissues develop and are maintained, and also how these processes are corrupted during oncogenic transformation.

**Properties that define stem cells**

Somatic stem cells are capable of specialized "self-renewing" mitotic divisions in which one or both of the daughter cells are faithful reproductions of the parent stem cell. To put it another way, an adult stem cell is defined as a somatic cell that can undergo extensive cell division and can also give rise to other stem cells or to cells that eventually differentiate to form specialized cells. Therefore, a tissue stem cell must possess three qualities to perform its natural function: self-renewal (e.g. be able to produce more stem cells), differentiation and homeostatic maintenance of the stem cell compartment. The special property of self-renewal is perhaps a stem cell's most important function, and the key to its functional exploitation through experimentation. This type of replication serves to expand the stem cell self-renewal programs in response to systemic or local signals, triggering massive proliferation of downstream, further differentiated progenitor cells. The ability to self-renew enables maintenance of a tissue specific undifferentiated pool of cells in the organ or tissue that it supports. Here, the difference between stem and

progenitor cells can be distinguished. Stem cells produce all the cell types required for a particular tissue to carry out its function for the lifetime of the organism. Immediate downstream progenitor cells have similar differentiation potential as stem cells, but they lack self-renewal and therefore support only transient tissue maintenance or regeneration. Progenitor cells are also found in greater numbers than stem cells, indicating they are an intermediate population undergoing mitosis, another quality which may be exploited to segregate them from their stem cell parents. The ability to prospectively isolate hematopoietic stem cells (HSCs) played a central role in the identification of regulators of stem cell self-renewal[2-5].

The second function of a stem cell is to differentiate into the highly specialized cells of an individual tissue. The process of differentiation often inversely correlates with a cell's ability to divide[6]. The idea that stem cells may be responsible for producing multiple differentiated cells has been around for many years. Numerous *in vitro* and *in vivo* assays have been designed to test the developmental potential of putative stem and progenitor cells. In blood, the most well defined stem cell system, stem cells produce transient amplifying progenitor cells that rapidly proliferate for a short time before terminally differentiating into very specialized cells that are quiescent or die during normal tissue maintenance or damage[7]. As cells progress towards maturity and specialization, their ability to divide is coordinately reduced, resulting in a terminally differentiated cell with highly specialized functions. These terminally differentiated cells perform their normal function, cycle into quiescence, or initiate an apoptotic transcriptional program. Traditionally, stem cells were thought to produce multiple cell types, but recent work has shown there is no such requirement. The assumption was that

if a primitive cell gave rise to a single type of differentiated cell, then it would be a differentiation-restricted progenitor. For example, B and T-cell progenitors are now thought to be unipotent stem cell populations. They fit the criteria for somatic stem cells in that they are long lived and give rise to differentiated progeny. Studies on lymphoma support the notion that progenitor cells can be thought of as stem cell populations[8].

A third, perhaps overlooked, characteristic of stem cells is the homeostatic regulation of the stem cell compartment. Given the power of tissue stem cells in the context of their tissue environments, stringent regulatory programs govern how many stem cells are present and actively producing progeny at any given time. Perturbations in stem cell or progenitor compartments may easily progress down an oncogenic path, producing tumors that resemble their tissues of origin[9]. This hierarchical concept of cancer is known as the cancer stem cell hypothesis. This is in contrast with the stochastic model of oncogenesis which states that most cells have an equal ability to be transformed, with tumorigenic potential depending on cell cycle context and genomic instability.

In summary, the human body contains many tissues that contain stem cell populations. Somatic stem cells are defined by three major properties. First, stem cells can self-renew, giving rise to other stem cells. Second, stem cells differentiate into mature cells specialized to carry out the numerous functions a tissue in the body performs. Third, the stem cell number in any tissue is stringently regulated so as not to adversely affect the tissue it supports. Although significant advances have been made to understand what characteristics many somatic stem cells share, the underlying mechanisms that govern these processes remain poorly understood.

**Murine mammary development**

Mammary development in mouse proceeds through an ordered series of growth and differentiation[10]. Around embryonic day 11 (E11) development begins with the formation of five mammary placodes along each of two milk lines. About one day later, these placodes grow into epithelial buds that are distinguishable from the surrounding epidermis. At E15.5, each of the buds proliferates to form a sprout and invades into the underlying fat pad. Each of these sprouts forms a lumen, whose opening is on the surface of the skin where epidermal invagination creates the nipple. By E18.5, the sprouts have developed into small mammary glands[11]. This developmental program is abolished at parturition, at which point the gland grows isometrically until puberty. At the onset of puberty, when mice are between 3-4 weeks of age, the mammary gland goes through a period of rapid expansion and remodeling. Development of the mammary gland is driven by terminal end buds (TEBs). These highly proliferative structures are made up of a single layer of undifferentiated cap cells and multiple inner layers of body cells[12]. The body cells that are closest to the cap cell layer are highly proliferative but body cells that are further away from the cap cells undergo apoptosis and form the hollow lumen of the developing duct[13]. TEBs bifurcate during proliferation, creating the branching pattern typical of mammary epithelium, until they reach the boundaries of the fat pad around 6-7 weeks of age. Lateral branching (a.k.a. side branching or alveolar sprouting), which is different than TEB bifurcation, occurs along the subtending ducts and results in controlled invasion into the fat pad. Lateral branching is dependent on recurrent estrus cycles and pregnancy, which promote differentiation of alveolar progenitors[14].

Puberty is evidenced by the rapid growth of the ductal epithelium into the surrounding stromal fat pad, and is modulated by many hormones. The proliferating mammary cells of the TEBs and ducts do not express steroid hormone receptors but the steroid receptor-positive cells are a separate population often found in close proximity. Presumably, the effect of hormones on TEB and ductal proliferation is indirect or paracrine as locally produced growth factors mediate the effect of the major mammogens.Some of these factors are direct target genes of the steroid hormone receptors. Importantly, depletion of these local factors results in phenocopies of knockout steroid hormone mouse models. The maturation of the hypothalamo-pituitary-gonadal axis initiates puberty by increasing the systemic level of gonadotrophins and promoting ovarian secretions of estrogen and progesterone. Estrogen is the dominant mammogen and the ovarian steroid that causes allometric growth of mammary epithelium[15]. Progesterone is required for tertiary lateral branching in the virgin gland[16].

In addition, there are many hormones that regulate mammary development during puberty and cause alveolar differentiation. Growth hormone, through Insulin-like growth factor 1 (IGF-1), promotes ductal growth and lateral branching of the mammary tree at puberty at the interface between the epithelium and stroma through an estrogen dependent mechanism[17,18]. Although growth hormone is a systemic hormone, IGF-1 is produced locally in the mammary gland. The vitamin D3 receptor antagonizes the proliferative signal from steroid hormones in the developing pubertal mammary gland[19]. Epidermal growth factor receptor (EGFR) is essential for pubertal ductal outgrowth and branching morphogenesis, but not alveolar development, and is expressed

by the stroma[20,21]. ErbB2 mediates timing of pubertal ductal elongation and fat pad invasion, but also does not affect lobuloalveolar development[22]. The EGFR ligand amphiregulin is expressed in the epithelial cells of the ducts and TEBs during pregnancy, and promotes ductal outgrowth at puberty. Prolactin is an important hormone that promotes lobuloalveolar development[23]. Oxytocin promotes differentiation and contractile activity in the mammary gland[24]. Many other hormones and molecules regulate mammary development, which have been thoroughly described elsewhere[25].

The adult glandular system is composed of three cell types (Figure 1.1). Myoepithelial cells are smooth muscle like epithelial cells that serve to contract and squeeze milk through the lumen of the ductal system until it is secreted from the nipples. These cells share characteristics with both classical epithelium and muscle cells, and thus are named accordingly. They form a single layer around mature ducts, bordering the basement membrane which separates the stroma from the ductal parenchyma. Luminal epithelial cells line the inside of the ducts as cuboidal cells. These cells are responsible for milk secretion into the ductal lumen; milk-secreting alveolar cells are thought to be derived from luminal cells. Luminal epithelium and myoepithelium are terminally differentiated cells that are quiescent, only displaying turnover correlating to stochastic loss of cells[26]. There are layers of periductal cells that surround the ductal tree, although their exact origin is still unknown[27].

**Evidence for stem cells in the murine mammary gland**

A number of the descriptive studies have suggested the existence of two stem cell populations in the ductal system[28-30]. Presumably, one stem cell population

is responsible for ductal elongation and the other for growth of secretory epithelium. In the mouse mammary gland, the stroma develops from birth but the epithelium develops postnatally, making the system highly amenable to manipulation. The mammary rudiment stemming from the fourth nipple (that would normally supply the fourth fat pad with ducts) may be efficiently removed, or cleared, in weaning mice. This procedure leaves behind a stromal fat pad cleared of endogenous epithelium, which provides a site for transplantation of either cells or tissue pieces. The original technique was described by DeOme and colleagues, where they transplanted small pieces of mammary epithelium into the cleared fat pads of recipient mice and found the donor ducts could engraft[31]. The donor ductal outgrowths could be serially transplanted for up seven times before senescence. This was the first experimental *in vivo* evidence providing the existence of mammary stem cells. Since the pioneering experiments by DeOme and colleagues, transplantation of cells into the cleared fat pads of recipient mice has been used to measure engraftment, differentiation, and self-renewal. In addition, mating the recipient mice allows investigation of how pregnancy affects mammary remodeling and morphogenesis. This technique was used to identify multipotent alveolar progenitor cells by transplanting pieces of mammary epithelium from pregnant and non-pregnant mice[32]. Later *in vivo* studies with single cell suspensions performed in limiting dilutions gave credence to a model of progenitor cells in both the ductal and alveolar compartments[33-35]. Surprisingly, the age of mice has less of an effect on stem cells than the number of mitotic divisions the epithelium undergoes, but alveolar cell maturation is unaffected by either chronology or division competency[36-38].

Many microscopic and ultra-structural studies have been undertaken to identify the various cell types that make up the gland[39-42]. The notion that the breast system contains a pool of undifferentiated stem cells is not a new one; in fact, scientists have known for decades that there are undifferentiated cells in mammary epithelium[43,44]. These "immature" cells were recognized by their lack of specialized intracellular organelles as well as indistinguishable apical and basal polarity. Common histological staining also failed to distinguish these cells among neighboring, more differentiated cells. They were named SLCs, or small electron-lucent cells, and were found near the basal lamina[45]. These cells are physically smaller than luminal cells, but were thought to be epithelial and not invading lymphocytes as they are found to be in mitotic pairing with cells that display epithelial characteristics. They were found at a frequency of 1-3% of the mammary epithelium, suggesting about $2\text{-}7x10^4$ of these cells per gland. Morphotypic analysis followed initial ultrastructural analyses. These studies used shape and type to describe the non-stained, or pale cells, and segregate them into two functionally different categories[29]. Furthermore, these cell types were present in the mammary glands from human, mouse and rat tissue.

Studies have also suggested the existence of a progenitor population that arises concomitant with alveolar differentiation. To date, there is no direct evidence that strongly supports the hypothesis that stem cells are located in TEBs, although many groups have suggested this possibility[43,46-48]. TEBs have no differentiated myoepithelial cells; in their place are a monolayer of basal cap cells. Cap cells are cuboidal in morphology and thought to be undifferentiated due to their lack of adhesion contacts with neighboring cells, diffuse chromatin arrangement, and lack of well defined

organelles. They are thought to be the progenitor myoepithelial cells, evidenced by marker staining and the acquisition of differentiated myofilamental structures Myoepithelial cells arise from the trailing edge of the cap cell layer[43]. In addition, cap cells also give rise to the differentiated luminal population, shown by marker staining which traced clusters of cap cells breaking away from their basal location and ingratiation into the luminal compartment of the TEB. This observation supports a hypothesis that stem cells may express a migratory program depending on their location within the mammary ductal system. Therefore, the model for ductal morphogenesis that has emerged has defined mammary stem cells as morphologically undifferentiated cells found at the leading edge of the invading mammary tree as well as along the basal lamina of the ducts. A recent study that analyzed the role of Bmi-1 in the maintenance of the ductal system[49] provided additional evidence for the presence of parity induced stem cells. *Bmi-1* is a gene that has been implicated in self-renewal of blood and neural stem cells[3,50,51]. The data showed a loss of both engraftment and self-renewal when donor *Bmi-1*$^{-/-}$ mammary epithelium was transplanted into syngeneic hosts that had their endogenous mammary epithelium removed. However, this effect was reversed when the mice were mated, suggesting activation of a pregnancy-related dormant stem cell population capable of multipotent differentiation that was able to bypass the Bmi-1 self-renewal signal.

Similar to the blood system, *in vitro* colony forming assays have proved valuable tools to determine the differentiation potential of isolated cell populations. There are two colony forming assays typically used in mammary biology[52]. The first is a co-culture system in which freshly dissociated cells are seeded onto an irradiated layer of fibroblasts

that serve as a feeder layer to promote epithelial-stromal interactions. At low densities and in combination with marker staining, this assay serves as both a clonal estimation of colony formation ability and differentiation potential. The second assay routinely used is semi-solid culturing where cells are seeded into Matrigel[53], an extracellular matrix milieu derived from EHS sarcoma cells *in vitro*. This material is rich in a wide variety of cytokines and growth factors; it has been extensively used to culture cells *in vitro* and promotes the engraftment[54] of tumor cells *in vivo*. As a colony forming system, the three-dimensional nature of the assay hinders accurate quantitation of colony formation, and staining the cultures for lineage specific antibodies also proves a technical challenge.

**Prospective isolation of murine mammary stem cells**

Given the limitations on purely descriptive analysis of mammary epithelial cells *ex vivo*, scientists turned to cell separation technologies to assess the functional attributes of different types of cells within the *in vivo* system. Using *in vivo* transplantation, numerous groups have attempted to assess the frequency of stem cells in single cell suspensions of murine mammary tissue[32,47,55,56]. These data have provided a highly variable estimation of stem cell numbers, ranging at the upper limit of less than 1 stem cell per 4900 cells to the lower limit of 1 stem cell in 100 cells. The large variability of stem cell frequency estimation comes from the variety of dissociation protocols and staining systems used. In addition, some groups mate recipient mice, potentially increasing the number of stem cells in the virgin mouse given the aforementioned observations of parity-induced multipotent stem cells and the powerful mitogenic effects of pregnancy related hormones. Even with these pro-growth signals, the experimental

manipulation of stem cells during in vivo studies probably underestimates the true stem cell frequency in a reported phenotypic population. Despite technical limitations, the transplant assay may also be used to estimate self-renewal of mammary stem cells. Reports have previously shown at least a 1000-fold expansion of stem cells upon serial transplantation. Two reports in 2006 identified phenotypic surface antigens that were used to enrich for the mammary stem cell population[55,56]. Taking advantage of the observation that any portion of the mature mammary gland was able to give rise to ductal outgrowths, these groups were able to use most of the fat pads from mature mice (except for the $5^{th}$ as it is technically difficult to resect) as source tissue. Although stem cells may be distributed throughout the mature gland, there is an increased number found in the lobules as opposed to the ducts, with the lowest amount found in alveolar lobules of lactating mice[48]. Both groups used FACS to prospectively isolate the self-renewing population from virgin mammary tissue, demonstrating these cells were capable of giving rise to mature luminal and myoepithelial cells. One group, led by John Stingl and Connie Eaves, proposed the mammary stem cell, or MRU (mammary repopulating unit), has a $CD24^{med}CD49f^{hi}Lin^{-}$ phenotype. The other group, led by Mark Shackleton and Jane Visvader, characterized the MaSC as having a $CD29^{hi}CD24^{+}Lin^{-}$ phenotype. Together they showed the stem cell phenotype was $CD24^{med}CD49^{hi}CD29^{hi}Sca\text{-}1^{low}Lineage^{-}$. Using this phenotype, the authors reported the frequency of stem cells as 1 per 60-90 cells. The authors described an 85% overlap between the $CD24^{med}CD49f^{hi}$ cells and the $CD29^{hi}CD24^{+}$ cells, showing that these phenotypes are ostensibly interchangeable. Unfortunately, there are no other reports that have been able to reproduce this frequency, although the phenotype of these cells has been confirmed. Both groups used a cocktail of

antibodies to mark "Lineage" made up of blood and endothelial cells. A number of markers have been used to delineate the mammary cellular hierarchy: CD24, CD49f, CD29, CD61, CD14, and Sca-1. Using combinations of these markers a progenitor population was identified as $CD24^{hi}CD29^{low}CD49f^{+}Lineage^{-}$. These cells had a dramatic reduction of *in vivo* engraftment ability compared to the MRU or MaSC, but could robustly produce colonies of both myoepithelial and luminal cells when placed into co-culture with irradiated NIH3T3 fibroblasts. They were named MaCFCs, or mammary colony forming cells, for this ability and they expressed luminal epithelial cytokeratin proteins suggesting they have a luminal phenotype.

Recent reports have further elucidated the mammary hierarchy by subdividing the MaCFC progenitor population into those cells that are luminal restricted and those that generate myoepithelial progeny. The luminal progenitor cells had the $CD24^{hi}CD49f^{+}CD29^{low}CD14^{+}CD61^{+}$ phenotype, and the myoepithelial progenitor population was $CD24^{med}CD49^{med}CD29^{hi}$. Luminal progenitors may be further fractionated into $Sca\text{-}1^{+}ER^{+}$ cells (ER for estrogen receptor) and $Sca\text{-}1^{-}$ cells that express milk proteins. Estrogen receptor was not expressed by mammary stem cells, but was expressed later in differentiation on luminal epithelial cells. Progenitor MaCFC $CD29^{low}CD61^{-}$ cells expressed ERα, and thus may represent terminally differentiated cells since the $CD29^{low}CD61^{+}$ fraction is enriched for multipotent progenitors. The mammary hierarchy, as it has been described to date, is described in Figure 1.2. These previous studies showed there is a threshold of the absolute numbers of donor cells required for ductal outgrowth production. This is probably an indication of the damage cells undergo during processing and exogenous manipulation in combination with the

homing of duct forming cells in the local environment for their optimal growth, as well as a reflection of enrichment for stem cells achieved by antibodies rather than a purified population. Regardless of those limitations, the transplant assay isthe standard to measure a specific population's ability to produce ductal epithelium. When used in combination with software such as L-Calc, this assay provided a powerful tool to measure the frequency of mammary progenitor and stem cells. As evidence of the clonality of this approach, mixtures of genetically marked cells are able to produce outgrowths that arise from one genotype. However, as mentioned earlier, there is a threshold of absolute numbers of cells needed for such engraftment due to the loss of cells caused by the experimental manipulation required for *in vivo* transplantation. Higher numbers of transplanted cells resulted in ducts that were polyclonal. A caveat of the transplantation assay thus far is the lack of pure populations of donor cells. Therefore, it is possible that there is a combination of cells that must be engrafted to achieve optimal engraftment efficiency. However, studies at the single cell level have shown outgrowths, lending credence to the concept that only a single stem cell is needed to produce the entire mammary ductal system. This brings back the concept of a microenvironment, where there is still open debate of the identity of niche cell(s) for the mammary stem cell. This area is poorly understood and investigated; a stem cell may use stromal cells or components such as specific ECM molecules as its niche or may produce daughter cells that serve that function. This debate is not limited to the mammary stem cell field. In fact, it has been extensively investigated in the blood system[57-65].

**Signaling pathways that regulate stem cells also play a role in cancer**

Stem cells are minority population, so the inherent difficulty of studying molecular pathways with small amounts of cells has thus far slowed progress dramatically. Even with these drawbacks, multiple pathways have been shown to be crucial in somatic stem cells (Table 1). These pathways are often corrupted in cancer, suggesting a thorough understanding of these molecular mechanisms may help to develop future therapeutics.

Polycomb genes were initially discovered in Drosophila to be repressors of the Homeobox genes, and this holds true in mammals as well. Polycomb genes are thought to repress their targets through chromatin modifications. *Bmi-1*, a gene that encodes a member of the Polycomb protein family, has been shown to be crucial to self-renewal of somatic stem cells[3,51]. *Bmi-1* plays a crucial role in the self-renewal of hematopoietic, neural and mammary stem cells. Polycomb genes such as *Bmi-1* (PCGF4) have shown to be necessary for self-renewal in hematopoietic stem cells (HSCs) by suppressing the *Ink4a/Arf* locus and the p53 pathway. The loss of neural and blood stem cells in *Bmi-1*$^{-/-}$ mice results from their inability to self-renew. In addition, *Bmi-1*$^{-/-}$ mammary epithelium is stunted in development upon transplantation. Growing evidence in blood and brain support the hypothesis that somatic stem cells may be more sensitive to senescence pathways than more differentiated progeny[3,51,66]. Some genes whose expression appears to be suppressed by Bmi-1 play a role in stem cell biology. Expression of *p16*$^{Ink4a}$, a cell cycle inhibitor, is elevated in *Bmi-1* mutant mice. Recent studies showed that *p16*$^{Ink4a}$ plays a role in stem cell senescence in the blood, brain, and pancreatic islet cells[67-69].

Other genes linked to stem cell maintenance may act at least in part through *Bmi-1*. For Example, Sonic hedgehog (*Shh*) was shown to expand human blood progenitors in immunocompromised mice[70]. In the brain, *Shh* appears to activate *Bmi-1*[71]. The Hedgehog signaling pathway is crucial to the embryonic development of skin, hair follicles and sebaceous glands[72] and involved in postnatal and adult brain development[73]. Interestingly, *Shh* and *Ihh* were found disensible for mammary morphogenesis, which may indicate funtional redundancy with another Hedgehog ligand. Mutation of *Shh* is known to cause Gorlin's syndrome, while activation of *Shh* has been implicated in both skin and brain carcinogenesis[74].

The Wnt/β-catenin pathway, which is associated with many types of cancer, also has been implicated in self-renewal[75]. Secreted Wnt ligands bind to Frizzled receptors and activate a cascade important in development. Progesterone is known to exert its effects on lateral branching in the mammary gland via Wnt4[76]. Wnt inhibitors have been shown to retard hematopoietic reconstitution *in vivo*. Wnt signaling increases *HoxB4* and *Notch-1* expression, both of which are implicated in the specification and/or self-renewal of hematopoietic stem cells [77]. In addition, Wnt/β-catenin pathway is involved in the maintenance of normal intestinal epithelial cells and implicated in regenerative responses during tissue repair[78,79].

The Notch signaling cascade is a transmembrane system widely shared by various animal cells for regulating embryonic development and adult maintenance of homeostasis. Well conserved from nematode to humans, the Notch signaling pathway in mammalians consists of four Notch receptors (Notch 1 to Notch 4) and four Notch ligands that physically bind to the Notch receptors. The roles of Notch signaling pathway

on stem cells and early progenitor cells are demonstrated in a number of studies of gene-modified animal models to influence HSC and melanocyte generation, CNS and vasculature development, organogenesis during embryogenesis; and to influence adult hematopoietic and immune systems, intestinal mucosal systems, skeletal muscle, skin and hair systems[80]. The Notch signaling regulates neural stem cell expansion *in vivo* and *in vitro*[81]. Mutations or aberrant activation of Notch signaling pathway is known to cause T-ALL, indicating its contribution to both normal development and carcinogenesis. A recent study showed knockdown of the Notch effector Cbf-1 increased mammary stem cell activity and also drove the formation of aberrant TEBs[82]. Importantly, constitutive Notch signaling resulted in expansion of luminal cells, leading to tumor formation.

The Hox gene family has been shown to be important in normal and malignant hematopoiesis. Several translocations in human leukemia involve Hox genes such as HoxA9 in AML [83]. In mouse models, aberrant expression of Hox genes has been shown to affect the proliferation and differentiation of HSCs[84]. When overexpressed in mouse bone marrow cells, several members of the Hox family[85,86] result in expansion of HSCs and myeloid precursor cells. Hoxc-6 transcripts are present in pubertal mammary tissue but downregulated during pregnancy, and Hox-1 is upregulated in breast cancers[87].

In mice, loss of expression of Pten, a gene that is implicated in the maintenance of hematopoietic stem cells and neural stem cells, and commonly mutated in many types human cancer [88], was recently shown to drive aberrant self-renewal of HSCs and eventually leads to leukemia in the mice [89]. Pten normally functions by limiting activity of inositol triphosphate signaling. Loss of Pten, which results constitutive

activation of the inositol triphosphate signaling pathway, initially promoted adult hematopoietic stem cell proliferation, but after a brief period of time there was subsequent exhaustion of the HSC pool, suggesting Pten is important in HSC maintenance[90]. Expression of Wnt-1 in mammary glands of transgenic mice resulted in tumors that had loss of Pten, suggesting it also has a role in control proliferation of mammary progenitor cells[91].

Most ABCG family of transporters play a significant role in the ATP dependent efflux transport of cholesterol. These proteins are mainly responsible for the side population phenotype which is defined by the extrusion of the DNA binding dye Hoechst 33342. The basis of these proteins as markers for self-renewing stem populations comes from the idea that long lived cell populations are constantly bombarded with genotoxic chemicals, and thus it is more efficient to efflux these chemicals from the cell rather than process them through cytoplasmic degradation machinery. When both blood and solid organs are dissociated and subjected to staining with dyes that are known to be effluxed, the minor population of cells that upregulate the transporters are revealed. In some but not all normal and cancer cells, it is thought that the stem cells contain some self-renewal and differentiation potential *in vitro[92]*. Although useful in other stem cell systems, murine mammary populations could not be enriched based on side population status[56].

Telomere shortening has been implicated in replicative senescence, chromosome instability and arrest of the cell cycle[93]. Since stem cells live for long periods of time in part by activating telomerase activity, this enzyme may play an important role in cancer stem cell biology as well. Telomerase, an enzyme that adds terminal repeats to the end of telomeres as a clock mechanism, was found to be expressed at high levels in normal self-

renewing populations in the blood[94,95] and tumor populations[96]. The telomerase protein has been shown to be crucial to the extension of somatic human cells[97] and is transcript expression upregulated in tumor cell mRNA[98,99]. hTERT, the catalytic subunit of telomerase, activation was found to be involved in tumorigenic transformation of cultured human skeletal muscle myoblasts and mammary cells [100].

| Pathway | Normal stem cells | Cancer |
|---|---|---|
| Bmi-1 | Bmi-1 is required for self-renewal of hematopoietic and neural stem cells | Bmi-1 is upregulated in AML and overexpressed in medulloblastoma |
| | Bmi-1 downregulates the Ink4a/Arf locus | Overexpressed Bmi-1 and cell proliferation induce self-renewal of leukemic stem cells |
| Shh | Involved in the maintenance of hematopoietic stem cells and expansion of progenitors | Activation of SHH is implicated in skin and brain carcinogenesis, including basal cell carcinoma of skin and medulloblastoma |
| | Crucial in embryonic development of skin, hair follicle, and sebaceous gland | Mutation of SHH causes Gorlin's syndrome |
| | Involved in postnatal and adult brain development | |
| Wnt/β-catenin | Involved in the maintenance and self-renewal of hematopoietic stem cells and progenitor cells | Overexpression of WNT is seen in many human cancers |
| | Regulates the maintenance of normal intestinal epithelial cells | Accumulation of β-catenin is associated with breast cancer, melanoma, sarcoma, myeloid leukemia, multiple myeloma, and brain tumors |
| | Implicated in regenerative responses during tissue repair | Mutations in β-catenin are found in endometrial carcinomas, prostate carcinoma, and hepatocellular carcinomas |
| | | Mutations of both β-catenin and APC genes are common in colorectal cancer |
| Notch | Mediates the self-renewal of hematopoietic and neural stem cells | Mutations or aberrrant activation of Notch1 are known to cause T-ALL in human and mouse |
| | Activates Notch target genes involved in T cell differentiation and self-renewal | |
| *Hox* family | Involved in the self-renewal of hematopoietic stem cells and the proliferation and differentiation of precursor cells | Overexpression of HOXA9 is found in AML patients with poor prognosis |
| | | Overexpression of HOX11 is described in T-ALL with chromosome translocations |
| | | Hoxb3, Hoxb8, and Hoxa10 are associated with leukemogenesis in a mouse model |
| Pten | Implicated in the maintenance of hematopoietic stem cells and neural stem cells | Loss of PTEN leads to the formation of a variety of tumors, including myeloproliferative disease, and the emergence of transplantable leukemia |
| | | Mutation and/or LOH cause glioblastoma multiforme, prostate carcinoma, and endometrial carcinoma |
| Efflux transporters | Marker proteins are found in self-renewing stem cells, such as ABCG family proteins, responsible for the side-population phenotype | Upregulated ABCG2, ABCB1, and CEACAM6 are found in cancer cells from the gastrointestinal system |
| | | Upregulated ABCG is implicated in broad-spectrum chemoresistance of cancer cells, such as AML cells |
| Telomerase | Expressed at a high level in normal self-renewing populations in the blood | Expressed at a high level in tumor cell populations with upregulated mRNA expression |
| | | hTERT is involved in tumorigenic transformation |
| | | Upregulated telomerase activity is found in glioblastoma |

[a]Abbreviations used: AML: acute myeloid leukemia; APC: adenomatous polyposis coli; hTERT: human telomerase reverse transcriptase; LOH: loss of heterogeneity; PTEN: phosphatase and tensin homolog deleted from chromosome 10; Shh/SHH: sonic hedgehog; T-ALL: T cell acute lymphoblastic leukemia; WNT: Wingless-Int.

**Table 1.1: Stem cell pathways that are also known to be corrupted in cancer.**

Many pathways have been implicated in stem cell self-renewal and differentiation. Not surprisingly, a number of these are found dysregulated in cancer. Adapted from Lobo, N. A., Y. Shimono, et al. Annu Rev Cell Dev Biol 23:675-99.

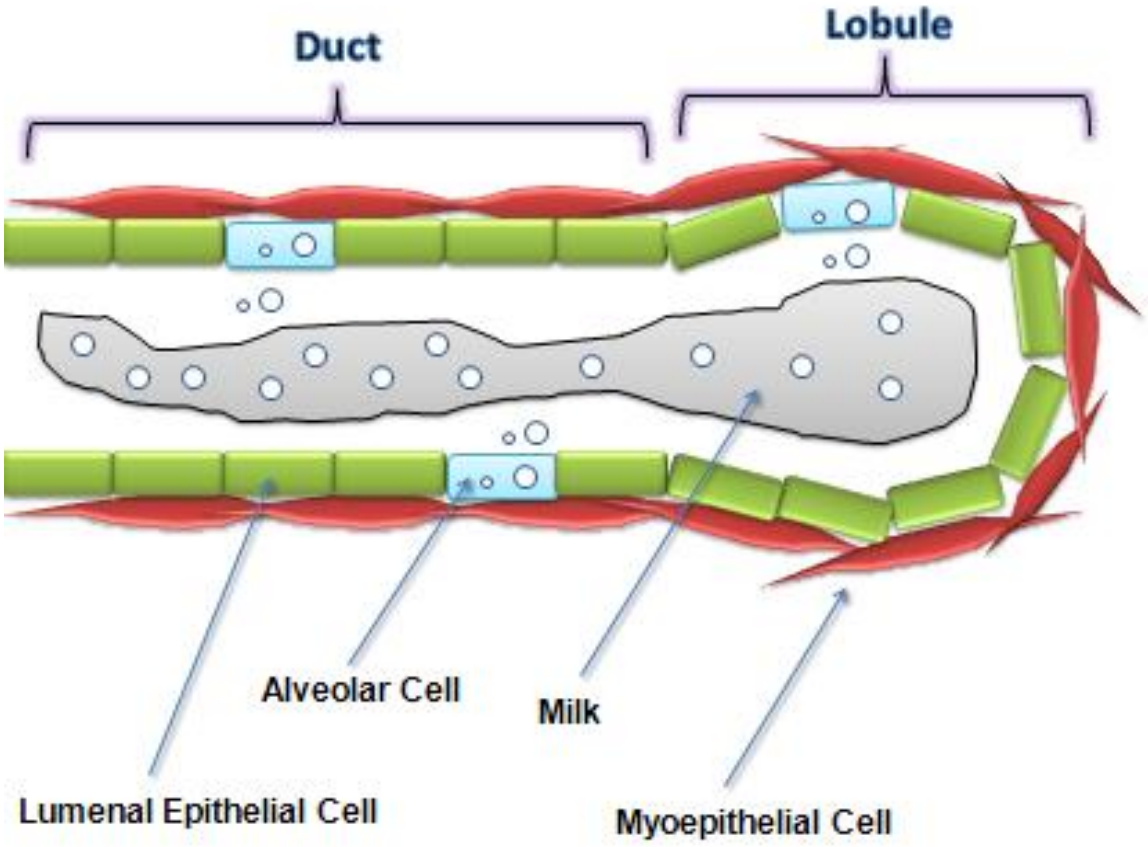



**Figure 1.1: Cartoon depiction of a mature terminal lobule.**

The mature mammary gland is composed of multiple cell type that carries out specific functions. Luminal epithelial cells are responsible for structure and secretion. Alveolar cells are the milk producing cells. Myoepithelial cells are contractile and force the milk through the lumen of the system until it is dispersed through the nipple.

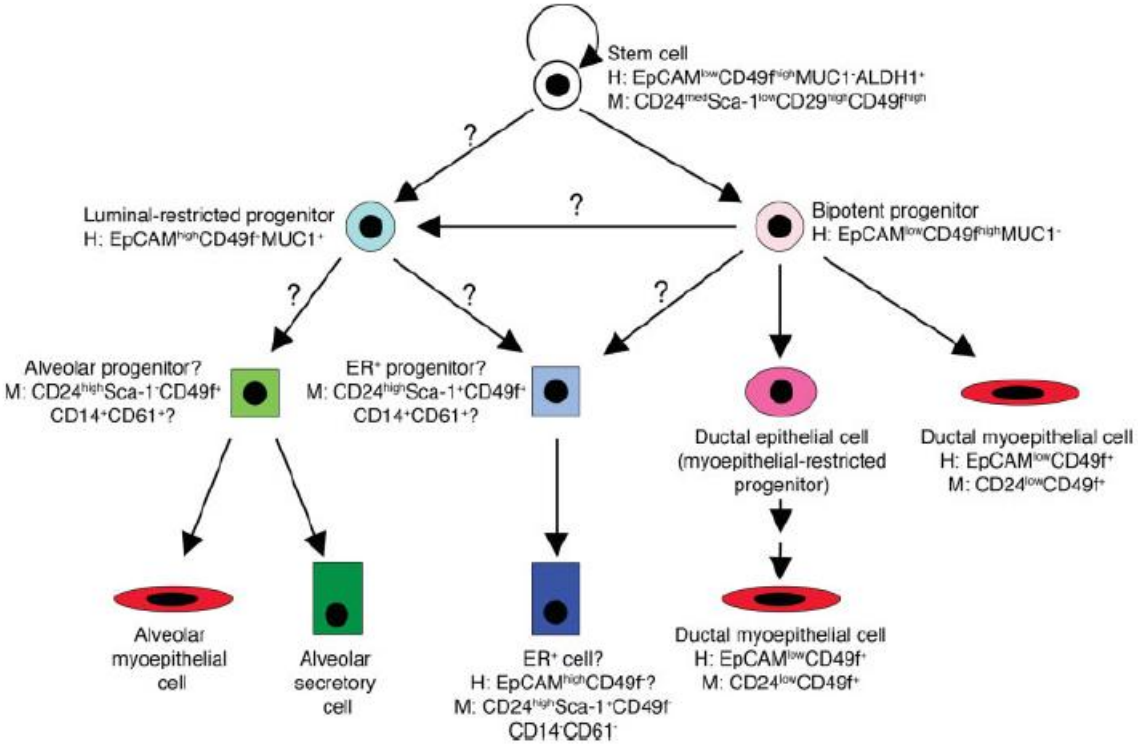



**Figure 1.2: Mammary cellular hierarchy of human and mouse.**

Present in human (H) and mouse (M). Many cell types have been described in the mammary system. However, intermediate early progenitors still remain poorly defined. Adapted from Stingl, J Pathol 217:229-241.

**BIBLIOGRAPHY**


1. McClay DR: **Gastrulation**. *Curr Opin Genet Dev* 1991, **1**:191-195.
2. Park IK, He Y, Lin F, Laerum OD, Tian Q, Bumgarner R, Klug CA, Li K, Kuhr C, Doyle MJ, et al.: **Differential gene expression profiling of adult murine hematopoietic stem cells**. *Blood* 2002, **99**:488-498.
3. Park IK, Qian D, Kiel M, Becker MW, Pihalja M, Weissman IL, Morrison SJ, Clarke MF: **Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells**. *Nature* 2003, **423**:302-305.
4. Cheng T, Rodrigues N, Shen H, Yang Y-G, Dombkowski D, Sykes M, Scadden DT: **Hematopoietic Stem Cell Quiescence Maintained by p21cip1/waf1.** *Science* 2000, **287**:1804-1808.
5. Calvi LM, Adams GB, Weibrecht KW, Weber JM, Olson DP, Knight MC, Martin RP, Schipani E, Divieti P, Bringhurst FR, et al.: **Osteoblastic cells regulate the haematopoietic stem cell niche**. *Nature* 2003, **425**:841-846.
6. Clarke MF, Fuller M: **Stem cells and cancer: two faces of eve**. *Cell* 2006, **124**:1111-1115.
7. Reya T, Morrison SJ, Clarke MF, Weissman IL: **Stem cells, cancer, and cancer stem cells**. *Nature* 2001, **414**:105-111.
8. Bernt KM, Armstrong SA: **Leukemia stem cells and human acute lymphoblastic leukemia**. *Semin Hematol* 2009, **46**:33-38.
9. Lobo NA, Shimono Y, Qian D, Clarke MF: **The biology of cancer stem cells**. *Annu Rev Cell Dev Biol* 2007, **23**:675-699.
10. Richert MM, Schwertfeger KL, Ryder JW, Anderson SM: **An atlas of mouse mammary gland development**. *J Mammary Gland Biol Neoplasia* 2000, **5**:227-241.
11. Veltmaat JM, Mailleux AA, Thiery JP, Bellusci S: **Mouse embryonic mammogenesis as a model for the molecular regulation of pattern formation**. *Differentiation* 2003, **71**:1-17.
12. Ball SM: **The development of the terminal end bud in the prepubertal-pubertal mouse mammary gland**. *Anat Rec* 1998, **250**:459-464.
13. Humphreys RC, Krajewska M, Krnacik S, Jaeger R, Weiher H, Krajewski S, Reed JC, Rosen JM: **Apoptosis in the terminal endbud of the murine mammary gland: a mechanism of ductal morphogenesis**. *Development* 1996, **122**:4013-4022.
14. Brisken C: **Hormonal control of alveolar development and its implications for breast carcinogenesis**. *J Mammary Gland Biol Neoplasia* 2002, **7**:39-48.
15. Hovey RC, Trott JF, Vonderhaar BK: **Establishing a framework for the functional mammary gland: from endocrinology to morphology**. *J Mammary Gland Biol Neoplasia* 2002, **7**:17-38.
16. Brisken C, Park S, Vass T, Lydon JP, O'Malley BW, Weinberg RA: **A paracrine role for the epithelial progesterone receptor in mammary gland development**. *Proc Natl Acad Sci U S A* 1998, **95**:5076-5081.
17. Kleinberg DL: **Early mammary development: growth hormone and IGF-1**. *J Mammary Gland Biol Neoplasia* 1997, **2**:49-57.

18. Gallego MI, Binart N, Robinson GW, Okagaki R, Coschigano KT, Perry J, Kopchick JJ, Oka T, Kelly PA, Hennighausen L: **Prolactin, growth hormone, and epidermal growth factor activate Stat5 in different compartments of mammary tissue and exert different and overlapping developmental effects**. *Dev Biol* 2001, **229**:163-175.
19. Zinser G, Packman K, Welsh J: **Vitamin D(3) receptor ablation alters mammary gland morphogenesis**. *Development* 2002, **129**:3067-3076.
20. Xie W, Paterson AJ, Chin E, Nabell LM, Kudlow JE: **Targeted expression of a dominant negative epidermal growth factor receptor in the mammary gland of transgenic mice inhibits pubertal mammary duct development**. *Mol Endocrinol* 1997, **11**:1766-1781.
21. Wiesen JF, Young P, Werb Z, Cunha GR: **Signaling through the stromal epidermal growth factor receptor is necessary for mammary ductal development**. *Development* 1999, **126**:335-344.
22. Stern DF: **ErbBs in mammary development**. *Exp Cell Res* 2003, **284**:89-98.
23. Naylor MJ, Lockefeer JA, Horseman ND, Ormandy CJ: **Prolactin regulates mammary epithelial cell proliferation via autocrine/paracrine mechanism**. *Endocrine* 2003, **20**:111-114.
24. Sapino A, Macri L, Tonda L, Bussolati G: **Oxytocin enhances myoepithelial cell differentiation and proliferation in the mouse mammary gland**. *Endocrinology* 1993, **133**:838-842.
25. Howlin J, McBryan J, Martin F: **Pubertal mammary gland development: insights from mouse models**. *J Mammary Gland Biol Neoplasia* 2006, **11**:283-297.
26. Villadsen R: **In search of a stem cell hierarchy in the human breast and its relevance to breast cancer evolution**. *APMIS* 2005, **113**:903-921.
27. Silberstein GB: **Tumour-stromal interactions. Role of the stroma in mammary development**. *Breast Cancer Res* 2001, **3**:218-223.
28. Smith GH, Mehrel T, Roop DR: **Differential keratin gene expression in developing, differentiating, preneoplastic, and neoplastic mouse mammary epithelium**. *Cell Growth Differ* 1990, **1**:161-170.
29. Smith GH, Medina D: **A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland**. *J Cell Sci* 1988, **90 ( Pt 1)**:173-183.
30. Chepko G, Smith GH: **Mammary epithelial stem cells: our current understanding**. *J Mammary Gland Biol Neoplasia* 1999, **4**:35-52.
31. Deome KB, Faulkin LJ, Jr., Bern HA, Blair PB: **Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice**. *Cancer Res* 1959, **19**:515-520.
32. Kordon EC, Smith GH: **An entire functional mammary gland may comprise the progeny from a single cell**. *Development* 1998, **125**:1921-1930.
33. Booth BW, Boulanger CA, Smith GH: **Selective segregation of DNA strands persists in long-label-retaining mammary cells during pregnancy**. *Breast Cancer Res* 2008, **10**:R90.
34. Booth BW, Boulanger CA, Smith GH: **Alveolar progenitor cells develop in mouse mammary glands independent of pregnancy and lactation**. *J Cell Physiol* 2007, **212**:729-736.

35. Boulanger CA, Wagner KU, Smith GH: **Parity-induced mouse mammary epithelial cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression**. *Oncogene* 2005, **24**:552-560.
36. Bocchinfuso WP, Lindzey JK, Hewitt SC, Clark JA, Myers PH, Cooper R, Korach KS: **Induction of mammary gland development in estrogen receptor-alpha knockout mice**. *Endocrinology* 2000, **141**:2982-2994.
37. Militzer K, Schwalenstocker H: **Postnatal and postpartal morphology of the mammary gland in nude mice**. *J Exp Anim Sci* 1996, **38**:1-12.
38. Greenman DL, Highman B, Chen JJ, Schieferstein GJ, Norvell MJ: **Influence of age on induction of mammary tumors by diethylstilbestrol in C3H/HeN mice with low murine mammary tumor virus titer**. *J Natl Cancer Inst* 1986, **77**:891-898.
39. Han J, Cao S, Jin H, Liu Y, Wang M, Song J, Li N: **Localization of putative stem cells and four cell populations with different differentiation degree in mouse mammary anlagen**. *Histochem Cell Biol* 2006, **126**:35-43.
40. Forster C, Makela S, Warri A, Kietz S, Becker D, Hultenby K, Warner M, Gustafsson JA: **Involvement of estrogen receptor beta in terminal differentiation of mammary gland epithelium**. *Proc Natl Acad Sci U S A* 2002, **99**:15578-15583.
41. Hogg NA, Harrison CJ, Tickle C: **Lumen formation in the developing mouse mammary gland**. *J Embryol Exp Morphol* 1983, **73**:39-57.
42. Rosano TG, Lee SK, Jones DH: **Developmental changes in mitochondria during the transition into lactation in the mouse mammary gland. II. Membrane marker enzymes and membrane ultrastructure**. *J Cell Biol* 1976, **69**:581-588.
43. Williams JM, Daniel CW: **Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis**. *Dev Biol* 1983, **97**:274-290.
44. Kordon EC: **MMTV-induced pregnancy-dependent mammary tumors : early history and new perspectives**. *J Mammary Gland Biol Neoplasia* 2008, **13**:289-297.
45. Ferguson DJ: **Ultrastructural characterisation of the proliferative (stem?) cells within the parenchyma of the normal "resting" breast**. *Virchows Arch A Pathol Anat Histopathol* 1985, **407**:379-385.
46. Kenney NJ, Smith GH, Lawrence E, Barrett JC, Salomon DS: **Identification of Stem Cell Units in the Terminal End Bud and Duct of the Mouse Mammary Gland**. *J Biomed Biotechnol* 2001, **1**:133-143.
47. Smith GH: **Label-retaining epithelial cells in mouse mammary gland divide asymmetrically and retain their template DNA strands**. *Development* 2005, **132**:681-687.
48. Sonnenberg A, Daams H, Van der Valk MA, Hilkens J, Hilgers J: **Development of mouse mammary gland: identification of stages in differentiation of luminal and myoepithelial cells using monoclonal antibodies and polyvalent antiserum against keratin**. *J Histochem Cytochem* 1986, **34**:1037-1046.
49. Pietersen AM, Evers B, Prasad AA, Tanger E, Cornelissen-Steijger P, Jonkers J, van Lohuizen M: **Bmi1 regulates stem cells and proliferation and differentiation of committed cells in mammary epithelium**. *Curr Biol* 2008, **18**:1094-1099.

50. Molofsky AV, He S, Bydon M, Morrison SJ, Pardal R: **Bmi-1 promotes neural stem cell self-renewal and neural development but not mouse growth and survival by repressing the p16Ink4a and p19Arf senescence pathways**. *Genes Dev* 2005, **19**:1432-1437.
51. Molofsky AV, Pardal R, Iwashita T, Park IK, Clarke MF, Morrison SJ: **Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation**. *Nature* 2003, **425**:962-967.
52. Stingl J: **Detection and analysis of mammary gland stem cells**. *J Pathol* 2009, **217**:229-241.
53. Kleinman HK, Martin GR: **Matrigel: basement membrane matrix with biological activity**. *Semin Cancer Biol* 2005, **15**:378-386.
54. Sterling-Levis K, White L: **The role of xenografting in pediatric brain tumor research with specific emphasis on medulloblastoma/primitive neuroectodermal tumors of childhood**. *In Vivo* 2003, **17**:329-342.
55. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, Wu L, Lindeman GJ, Visvader JE: **Generation of a functional mammary gland from a single stem cell**. *Nature* 2006, **439**:84-88.
56. Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, Li HI, Eaves CJ: **Purification and unique properties of mammary epithelial stem cells**. *Nature* 2006, **439**:993-997.
57. Can A: **Haematopoietic stem cells niches: interrelations between structure and function**. *Transfus Apher Sci* 2008, **38**:261-268.
58. Kubota Y, Takubo K, Suda T: **Bone marrow long label-retaining cells reside in the sinusoidal hypoxic niche**. *Biochem Biophys Res Commun* 2008, **366**:335-339.
59. Wagner W, Horn P, Bork S, Ho AD: **Aging of hematopoietic stem cells is regulated by the stem cell niche**. *Exp Gerontol* 2008, **43**:974-980.
60. Yahata T, Muguruma Y, Yumino S, Sheng Y, Uno T, Matsuzawa H, Ito M, Kato S, Hotta T, Ando K: **Quiescent human hematopoietic stem cells in the bone marrow niches organize the hierarchical structure of hematopoiesis**. *Stem Cells* 2008, **26**:3228-3236.
61. Kiel MJ, Acar M, Radice GL, Morrison SJ: **Hematopoietic stem cells do not depend on N-cadherin to regulate their maintenance**. *Cell Stem Cell* 2009, **4**:170-179.
62. Kiel MJ, Radice GL, Morrison SJ: **Lack of evidence that hematopoietic stem cells depend on N-cadherin-mediated adhesion to osteoblasts for their maintenance**. *Cell Stem Cell* 2007, **1**:204-217.
63. Kiel MJ, Morrison SJ: **Uncertainty in the niches that maintain haematopoietic stem cells**. *Nat Rev Immunol* 2008, **8**:290-301.
64. Kiel MJ, Morrison SJ: **Maintaining hematopoietic stem cells in the vascular niche**. *Immunity* 2006, **25**:862-864.
65. Morrison SJ, Wright DE, Cheshier SH, Weissman IL: **Hematopoietic stem cells: challenges to expectations**. *Curr Opin Immunol* 1997, **9**:216-221.
66. Ohtani N, Yamakoshi K, Takahashi A, Hara E: **The p16INK4a-RB pathway: molecular link between cellular senescence and tumor suppression**. *J Med Invest* 2004, **51**:146-153.

67. Janzen V, Forkert R, Fleming HE, Saito Y, Waring MT, Dombkowski DM, Cheng T, DePinho RA, Sharpless NE, Scadden DT: **Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16INK4a**. *Nature* 2006, **443**:421-426.
68. Krishnamurthy J, Ramsey MR, Ligon KL, Torrice C, Koh A, Bonner-Weir S, Sharpless NE: **p16INK4a induces an age-dependent decline in islet regenerative potential**. *Nature* 2006, **443**:453-457.
69. Molofsky AV, Slutsky SG, Joseph NM, He S, Pardal R, Krishnamurthy J, Sharpless NE, Morrison SJ: **Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing**. *Nature* 2006, **443**:448-452.
70. Bhardwaj G, Murdoch B, Wu D, Baker DP, Williams KP, Chadwick K, Ling LE, Karanu FN, Bhatia M: **Sonic hedgehog induces the proliferation of primitive human hematopoietic cells via BMP regulation**. *Nat Immunol* 2001, **2**:172-180.
71. Leung C, Lingbeek M, Shakhova O, Liu J, Tanger E, Saremaslani P, Van Lohuizen M, Marino S: **Bmi1 is essential for cerebellar development and is overexpressed in human medulloblastomas**. *Nature* 2004, **428**:337-341.
72. Athar M, Tang X, Lee JL, Kopelovich L, Kim AL: **Hedgehog signalling in skin development and cancer**. *Exp Dermatol* 2006, **15**:667-677.
73. Palma V, Lim DA, Dahmane N, Sanchez P, Brionne TC, Herzberg CD, Gitton Y, Carleton A, Alvarez-Buylla A, Ruiz i Altaba A: **Sonic hedgehog controls stem cell behavior in the postnatal and adult brain**. *Development* 2005, **132**:335-344.
74. Ruiz i Altaba A, Sanchez P, Dahmane N: **Gli and hedgehog in cancer: tumours, embryos and stem cells**. *Nat Rev Cancer* 2002, **2**:361-372.
75. Taipale J, Beachy PA: **The Hedgehog and Wnt signalling pathways in cancer**. *Nature* 2001, **411**:349-354.
76. Lamb R, Harrison H, Clarke RB: **Mammary development, carcinomas and progesterone: role of Wnt signalling**. *Ernst Schering Found Symp Proc* 2007:1-23.
77. Reya T, Duncan AW, Ailles L, Domen J, Scherer DC, Willert K, Hintz L, Nusse R, Weissman IL: **A role for Wnt signalling in self-renewal of haematopoietic stem cells**. *Nature* 2003, **423**:409-414.
78. de Lau W, Barker N, Clevers H: **WNT signaling in the normal intestine and colorectal cancer**. *Front Biosci* 2007, **12**:471-491.
79. Reguart N, He B, Taron M, You L, Jablons DM, Rosell R: **The role of Wnt signaling in cancer and stem cells**. *Fut Oncol* 2005, **1**:787-797.
80. Chiba S: **Notch signaling in stem cell systems**. *Stem Cells* 2006, **24**:2437-2447.
81. Androutsellis-Theotokis A, Leker RR, Soldner F, Hoeppner DJ, Ravin R, Poser SW, Rueger MA, Bae SK, Kittappa R, McKay RD: **Notch signalling regulates stem cell numbers in vitro and in vivo**. *Nature* 2006, **442**:823-826.
82. Bouras T, Pal B, Vaillant F, Harburg G, Asselin-Labat ML, Oakes SR, Lindeman GJ, Visvader JE: **Notch signaling regulates mammary stem cell function and luminal cell-fate commitment**. *Cell Stem Cell* 2008, **3**:429-441.
83. Nakamura T, Largaespada DA, Lee MP, Johnson LA, Ohyashiki K, Toyama K, Chen SJ, Willman CL, Chen IM, Feinberg AP, et al.: **Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia.** *Nat Genet* 1996, **12**:154-158.

84. Lawrence HJ, Helgason CD, Sauvageau G, Fong S, Izon DJ, Humphries RK, Largman C: **Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis.** *Blood* 1997, **89**:1922-1930.
85. Kirito K, Fox N, Kaushansky K: **Thrombopoietin induces HOXA9 nuclear transport in immature hematopoietic cells: potential mechanism by which the hormone favorably affects hematopoietic stem cells**. *Mol Cell Biol* 2004, **24**:6751-6762.
86. Beslu N, Krosl J, Laurin M, Mayotte N, Humphries KR, Sauvageau G: **Molecular interactions involved in HOXB4-induced activation of HSC self-renewal**. *Blood* 2004, **104**:2307-2314.
87. Friedmann Y, Daniel CA, Strickland P, Daniel CW: **Hox genes in normal and neoplastic mouse mammary gland**. *Cancer Res* 1994, **54**:5981-5985.
88. Wei Q, Clarke L, Scheidenhelm DK, Qian B, Tong A, Sabha N, Karim Z, Bock NA, Reti R, Swoboda R, et al.: **High-grade glioma formation results from postnatal pten loss or mutant epidermal growth factor receptor expression in a transgenic mouse glioma model**. *Cancer Res* 2006, **66**:7429-7437.
89. Zhang J, Grindley JC, Yin T, Jayasinghe S, He XC, Ross JT, Haug JS, Rupp D, Porter-Westpfahl KS, Wiedemann LM, et al.: **PTEN maintains haematopoietic stem cells and acts in lineage choice and leukaemia prevention**. *Nature* 2006, **441**:518-522.
90. Yilmaz OH, Valdez R, Theisen BK, Guo W, Ferguson DO, Wu H, Morrison SJ: **Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells**. *Nature* 2006, **441**:475-482.
91. Li Y, Welm B, Podsypanina K, Huang S, Chamorro M, Zhang X, Rowlands T, Egeblad M, Cowin P, Werb Z, et al.: **Evidence that transgenes encoding components of the Wnt signaling pathway preferentially induce mammary cancers from progenitor cells**. *Proc Natl Acad Sci U S A* 2003, **100**:15853-15858.
92. Jonker JW, Freeman J, Bolscher E, Musters S, Alvi AJ, Titley I, Schinkel AH, Dale TC: **Contribution of the ABC transporters Bcrp1 and Mdr1a/1b to the side population phenotype in mammary gland and bone marrow of mice**. *Stem Cells* 2005, **23**:1059-1065.
93. Counter CM, Avilion AA, LeFeuvre CE, Stewart NG, Greider CW, Harley CB, Bacchetti S: **Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity**. *Embo J* 1992, **11**:1921-1929.
94. Morrison SJ, Prowse KR, Ho P, Weissman IL: **Telomerase activity in hematopoietic cells is associated with self-renewal potential**. *Immunity* 1996, **5**:207-216.
95. Allsopp RC, Morin GB, DePinho R, Harley CB, Weissman IL: **Telomerase is required to slow telomere shortening and extend replicative lifespan of HSCs during serial transplantation**. *Blood* 2003, **102**:517-520.
96. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL, Shay JW: **Specific association of human telomerase activity with immortal cells and cancer**. *Science* 1994, **266**:2011-2015.

97. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, Harley CB, Shay JW, Lichtsteiner S, Wright WE: **Extension of life-span by introduction of telomerase into normal human cells**. *Science* 1998, **279**:349-352.
98. Saeboe-Larssen S, Fossberg E, Gaudernack G: **Characterization of novel alternative splicing sites in human telomerase reverse transcriptase (hTERT): analysis of expression and mutual correlation in mRNA isoforms from normal and tumour tissues**. *BMC Mol Biol* 2006, **7**:26.
99. Xin ZT, Beauchamp AD, Calado RT, Bradford JW, Regal JA, Shenoy A, Liang Y, Lansdorp PM, Young NS, Ly H: **Functional characterization of natural telomerase mutations found in patients with hematological disorders**. *Blood* 2006.
100. Kendall SD, Linardic CM, Adam SJ, Counter CM: **A network of genetic events sufficient to convert normal human cells to a tumorigenic state**. *Cancer Res* 2005, **65**:9824-9828.

# CHAPTER 2

## THY-1 DISCRIMINATES BETWEEN MAMMARY STEM CELLS AND MULTIPOTENT PROGENITOR CELLS


### SUMMARY

Thy-1 is a GPI-anchored cell surface protein that has been found on stem cells of the blood and liver. Although Thy-1 is a marker of myoepithelial differentiation in the rat mammary gland, little is known about expression of Thy-1 in murine breast. We recently showed Thy-1 marks tumorigenic cells in the *MMTV-Wnt1* cancer model. Previous groups have shown the stem cell population in murine breast has the $CD24^{med}CD49f^{hi}$ (MRU) phenotype. Our data shows that Thy-1 is differentially expressed between mammary stem cells and progenitor cells. $Thy\text{-}1^{+}CD24^{med}CD49f^{hi}$ cells are significantly enriched for ductal outgrowth forming cells *in vivo* limiting dilution transplantation. These cells are also enriched for self-renewing cells, as shown by their ability to produce secondary outgrowths. $Thy\text{-}1^{-}CD24^{med}CD49f^{hi}$ cells had reduced engraftment potential and severely diminished self-renewal. Our data suggests Thy-1 enriches for stem cells in the murine mammary gland. Furthermore, we propose the $Thy\text{-}1^{-}CD24^{med}CD49f^{hi}$ cells are mammary multipotent progenitor (MMPP) cells.

## INTRODUCTION

Recent reports presented evidence that a somatic stem cell population in the murine mammary system is responsible for both establishment and maintenance of the gland. Mammary stem cells are temporally regulated and respond to both systemic and local hormonal signals. The mammary stem cell population has been prospectively isolated using cell surface proteins in combination with FACS. Similar strategies have been used to identify other normal and cancer stem populations[1]. Increasingly, these protein "markers" are found to enrich for stem cells across multiple tissues suggesting all somatic stem cell populations share common traits, although the function of these markers may differ according to their tissue specific context.

Mammary glands develop from embryonic epidermis and begin as milk bud rudiments. These rudiments begin to proliferate and branch *in utero* forming milk buds, but become quiescent at parturition. Around week three of post-natal mouse development at the onset of puberty, secreted hormones stimulate growth of milk buds and promote invasion of the nascent branching system into the stromal fat pad, culminating in a mature gland made of ducts and terminal end bud (TEB) structures. Pregnancy is often characterized by alveolar cell differentiation followed by extensive lateral expansion of TEBs and subsequent milk secretion in response to lactogenic hormones. The ductal system in nulliparous mice has two main cell types: luminal epithelial cells that line the inside of ducts and myoepithelial cells that are muscle-like contractile epithelial cells that reside between the luminal cells and the basal lamina. Traditionally, cytokeratin

molecules have been used to identify the lineage of terminally differentiated cells, since stem cells acquire these markers. Myoepithelial cells expressed keratins (KRT) KRT5 and KRT14, as well as SMAα (smooth muscle actin alpha subunit). Mature luminal cells expressed KRT8 and KRT18. Mammary stem cells have recently been isolated and characterized by their surface phenotype $CD24^{med}CD49^{hi}$ (MRU) or $CD24^{+}CD29^{hi}$ [2,3]. Although different combinations of markers were used, other groups have shown that both strategies isolate similar populations with dramatic enrichment for mammary stem cells. Based on the marker combination presented by Stingl and colleagues, we sought to improve upon the current system by analyzing the expression and prospective enrichment for MRUs using Thy-1.

In mouse, Thy-1 is encoded on Chromosome 9 and has two allelic forms[4], Thy-1.1 and Thy-1.2, only one of which is expressed in a given mouse strain. It is a small GPI-anchored cell surface protein that has a single immunoglobulin domain, part of the larger Ly-6 protein family. Ly-6 genes have been linked to tumor cell adhesion, carcinogenesis, and cellular activation. Thy-1 proteins have been associated with T cell activation and proliferation, T-cell progenitor cell survival, cytokine and growth factor responses. Thy-1 was originally described as an alloantigen expressed on mouse thymocytes[5]. It has long been used as a hematopoietic progenitor marker, where its expression is conserved between rat, mouse and human systems[6-12]. Thy-1 has been shown to mark hepatic progenitor cells[13,14]}, murine mesenchymal stem cells[15] and neural cells[16,17].

Thy-1 has also been shown to be expressed in the normal mouse mammary gland and tumors[18,19]. In previous studies, Thy-1 was widely used a marker of myoepithelial

cells in the rat mammary gland[20,21] *In vitro*, rat mammary epithelial cells that expressed Thy-1 formed webbed and stellate colonies similar to myoepithelial morphology[22]. Thy-$1^+$ cells in the rat liver also have a myofibroblast/stellate phenotype[23].

Our group has prospectively isolated a cancer stem cell population from a subset of *MMTV-Wnt1* mammary tumors based on the Thy-$1^+$CD$24^+$ phenotype (Cho et al 2008). These cells shared characteristics of myoepithelial cells such as expression of the basal KRT5, KRT14 and KRT17. Interestingly, the MRU population described by Stingl and colleagues were also characterized by upregulation of these keratins. The non-tumorigenic cells in *MMTV-Wnt1* tumors shared a similar luminal epithelial keratin profile with mammary colony forming cells (MaCFC), presumably the progenitor population, which included upregulated KRT18 and KRT19. Based on these observations, we hypothesized Thy-1 would mark stem cells in the normal mammary system and these cells would have an enriched capacity for duct formation and self-renewal.

## MATERIALS AND METHODS

### Animals

All animals used in the study were C57Bl/6 and pCx-GFP mice that were maintained at the Stanford Animal Facility in accordance with the guidelines of both Institutional Animal Care Use Committees.

**Mammary gland dissociation**

6-10 weeks old C57Bl/6 mice were euthanized and all fat pads were surgically resected. Tissue was placed into L-15 media and minced with a razor blade into approximately 1-3mm pieces. Either 8 units of Blendzyme 4 or 1 ml 10X magnification Collagenase/Hyaluronidase solution and 100 Kunitz units of DNase was added and tissue was digested for 1.5 hrs with mechanical dissociation performed every 30 min. ACK lysis was the performed to remove red blood cells. Cells were collected by centrifugation at 1400 rpm for 5 min at 4°C, resuspended in 5 mls of pre-warmed Trypsin, and gently pipetted for 2 minutes. Cells were then spun again at 1400 rpm for 5 min at 4C, and resuspended in pre-warmed 2 ml of Dispase and 100 units of DNase. Cells were gently pipetted for 2 minutes to release cells and then 13 mls of HBSS+2% calf serum staining media was added. The dissociated cell solution was put through a 40uM mesh filter and the resulting cell solution counted by hemacytometer to determine both cell number and cell viability. For all experiments, cells were >99% viable. Pelleted cells were resuspended at a concentration of $1x10^7$ per ml and subjected to staining for flow cytometry.

**Staining protocol for FACS**

Cells were resuspended at a concentration of $1 \times 10^6$ cells per 100 µl of HBSS with 2% calf serum. Cells were blocked with rabbit or mouse IgG (10 µg/ml) and antibodies were added at appropriate dilutions determined from titration experiments. For the normal mammary stem cell experiments, antibodies included CD49f, CD31, CD45,

Ter119 (BD Pharmingen), Sca-1, CD24, Thy1.2 and CD140a (eBioscience). Cells were stained for 20 min on ice and washed with staining media. When biotinylated primary antibodies were used, cells were further stained with streptavidin-conjugated fluorophores and washed. Ultimately, cells were resuspended in staining media containing 7-aminoactinomycin D (1 µg/ml final concentration) or 4'-6-diamidino-2-phenylindole (DAPI, 1 µg/ml final concentration) to stain dead cells.

For all experiments, cells were analyzed and sorted using a FACSAria cell sorter (BD Bioscience). Side scatter and forward scatter profiles were used to eliminate debris and cell doublets. Dead cells were eliminated by excluding $DAPI^+$ cells, whereas contaminating mouse $CD45^+CD31^+Ter119^+CD140\alpha^+$ cells were eliminated by excluding cells labeled with the fluorophore used for the lineage antibody cocktail. In cell-sorting experiments, cell populations underwent two consecutive rounds of purification (double sorting) when the initial purity was not deemed high enough (>80%) and a sufficient number of cells were available.

**Whole mount of mammary glands**

Whole mounts stained with Carmine Alum were processed according to previous reports[24]. Briefly, inguinal fat pads were excised and spread on glass slides. After air-drying to secure the glands, the tissue was fixed in Carnoy's solution for >4hrs, then washed with 70% EtOH and distilled water. The tissues were then stained with Carmine Alum solution overnight, dehydrated through graded alcohols, cleared in Xylene, and mounted in Permount. For fluorescent tissue, inguinal glands were placed onto glass

microscope slides and pictures were taken immediately upon excision. Glands were then processed for Carmine Alum staining as described to verify duct formation.

**Immunohistochemistry**

For Thy-1 and CD24 staining, paraffin sections were first pre-treated by microwave for 10 minutes in citrate buffer, pH 6.0. Sections were the processed through the M.O.M kit (Vector Labs) according to manufacturer's instructions. Thy-1.1 was added at a dilution of 1:25. Rat secondary was used at 1:200 dilution. Sections were counterstained with Harris Hematoxylin for 4 seconds, then dehydrates and mounted.

***In vitro* co-culture assay**

Co-culture with NIH3T3 cells was performed as described[2]. NIH3T3 cells were maintained in DMEM supplemented with 10%HICS and PSA. Briefly, sub-confluent NIH3T3 fibroblasts were trypsinized and subjected to 10Gy of γ-radiation. Fibroblasts were then immediately plated into 24-well tissue culture plates at a concentration of 20,000 cells/well. Complete Mouse Epicult (Stem Cell Technologies) plus 5% FBS and PSA (Gibco) was added and cells were allowed to attach for 2 hours prior to addition of double sorted mammary cell populations. Primary cells were sorted directly from the flow cytometer into the plates to reduce experiment manipulation. After 24 hours, media in all wells was replaced with complete Mouse Epicult media plus PSA without serum to promote epithelial cell growth. Colonies were fixed after 7 days with ice cold methanol:acetone for 2 min and allowed to dry. Colonies were rehydrated with distilled

water, and Wright-Giemsa stain applied for 1 min. Colonies were then washed with water, allowed to air dry, and pictures taken.

### *In vivo* transplants

Sorted cell populations were collected in staining media. Cells were resuspended at the correct concentration in staining media and injected into the cleared fat pads of 21-28 days old syngeneic recipient C57Bl/6 mice. For all injections of 600 cells and below, cell counts were verified using either a nuclear staining count or GFP+ cell count. Cells were injected in either 10 or 5 ul volumes using a 25 ul Hamilton syringe. All transplants were allowed to grow for at least 5 weeks but not more than 10 weeks before analysis. In the case of secondary transplants, 1-2 mm pieces of tissue were transplanted into the cleared fat pads of syngeneic recipient mice.

## RESULTS

### Thy-1 is expressed on mammary epithelial cells

We initially investigated the expression of Thy-1 in freshly dissociated mammary cells using flow cytometry. Mammary fat pads from 6-week old female C57Bl/6 mice were processed into single cells and stained with antibodies against CD24, CD49f, Thy-1.2 along with a lineage cocktail consisting of CD45, CD31, Ter119 and CD140α (Lineage). All of the mammary epithelial cell analyses were performed using criteria consisting of DAPI exclusion to remove dead cells (data not shown). Mammary epithelial cells accounted for approximately 44.3% (n=8) of all viable cells (Figure 2.1A). We defined our cell populations based on the nomenclature of Stingl et al[2]. In this system

the MRUs, or mammary repopulating units, are analogous to the stem cell population defined by phenotype as $CD24^{med}CD49f^{hi}$. The MaCFCs, or further differentiated progenitor cells, were $CD24^{hi}CD49f^{med}$. The MYO, or differentiated myoepithelial cells, are $CD24^{lo}CD49f^{med}$. We defined differentiated luminal epithelial cells (EPI) based on the $CD24^{med}CD49f^{-}$ phenotype. Figure 2.1B shows the gates that we used for the MRU, MaCFC, MYO and EPI populations. These gates represent cell types that correlate with mammary stem cells, progenitors, and the two terminally differentiated populations, respectively. We analyzed mammary epithelial cells (Figure 2.1C) and found MRUs accounted for an average of 1.25% and MaCFC progenitors an average of 2.44% of $Lineage^{-}$ cells. The MYO and EPI cells were 4.14% and 5.99% of $Lineage^{-}$ cells, respectively. We did not explore the identity of the large amount of cells that fall outside of the described gates.

Cells were then subjected to flow cytometry analysis to determine how Thy-1 was expressed on dissociated cells. Figure 2.1D shows the staining profile of mammary epithelial cells based on the expression of Thy-1.2 and CD24. We observed large $Thy\text{-}1^{+}CD24^{-}$ and $Thy\text{-}1^{-}CD24^{+}$ populations, with a smaller amount of cells positive for both proteins. We next investigated the percentage of MRU and MaCFC that expressed Thy-1 and found about half of each population stained positive (Figure 2.1E). We also showed that about 10% of MYO cells expressed Thy-1 but only a few EPI cells express the protein. To indirectly measure the amount of Thy-1 on the surface of MRUs and MaCFCs, we measured the mean fluorescence intensity of Thy-1 staining. We observed more Thy-1 protein expression on MRUs than MaCFCs, suggesting Thy-1 expression decreases as stem cells differentiate into progenitors (Figure 2.1F). As Thy-1 has been

described as a myoepithelial marker[25,26], our data is consistent with other reports that suggest mammary stem cells may reside near the basement membrane of ducts, where myoepithelial cells are found.

We analyzed Sca-1 expression in mammary populations (Figure 2.2) as it was reported as another stem cell marker[27,28]. In the initial report, Sca-1 GFP reporter mice were used for transplantation and the authors found only slowly cycling, Sca-$1^{+}$ cells were able to produce mammary outgrowths. These results contrasted the recent report by Stingl et al. that showed the MRU fraction enriched for function stem cells expressed the basal cell markers keratin 14 and SMAα[2]. This report also demonstrated the stem cells were Sca-$1^{low}$. To distinguish between these opposing results, we stained mammary cells with a bright Sca-1 FITC antibody in combination with our other standard markers. Consistent with data presented by Stingl and colleagues, we observed approximately 10% of MRU and MYO cells expressed Sca-1. In stark contrast, almost all of the MaCFC progenitor cells and about 25% of EPI cells expressed Sca-1. MaCFCs have a luminal phenotype, so our results are also consistent with the original report that showed Sca-1 positive cells were only found in the luminal compartment of mature ducts.

Using flow cytometry, we were able to quantify the amount of Sca-1 protein on the surface of the cells by taking a measurement of the mean fluorescence intensity of each population. We used this metric although it indirectly correlates protein quantity because of difficulties of obtaining enough of any of these populations for western blot analysis. Although the percentage of Thy-$1^{+}$MRU and Thy-$1^{-}$MRU cells that expressed Sca-1 did not differ, we found the intensity of fluorescent Sca-1 molecules on the surface of MRUs was than any other population (Figure 2.2B), suggesting they are Sca-$1^{hi}$,

whereas the MaCFC progenitors were Sca-$1^{low}$. Even though these results were interesting, in vivo engraftment of the Sca-1 subpopulations from the Thy-$1^{+}$MRU and Thy-$1^{-}$MRU populations is needed to validate of Sca-1 is a stem cell enriching marker.

**Thy-1 localizes to basal cells of mature ductal epithelium**

To address the location of Thy-1 expressing cells in mature ducts, we performed immunohistochemistry on paraffin sections made from the mammary fat pads of 6-week old C57BL/6 female mice. We first stained with hematoxylin and eosin (H&E) to visualize the cellularity of ducts (Figure 2.3A). Basal cells appeared pink and luminal cells stained purple. Using our H&E staining for reference, we proceeded to stain sections for Thy-1 and CD24. We observed specific red Thy-1 staining near the basal lamina that separates the myoepithelium and stromal cells, along with diffuse staining in the periductal regions (Figure 2.3B). This staining pattern was similar to CD49f, which localizes to the basal cell layer of ducts[2,29]. We also looked at CD24 expression (Figure 2.3C), and observed diffuse staining throughout the luminal and basal layers. Although histology was not sensitive enough to detect qualitative differences in expression, the general staining pattern was consistent with our flow cytometry expression data that CD24 is expressed by both luminal and myoepithelial cells.

**Thy-$1^{+}$CD24$^{med}$CD49f$^{hi}$ mammary cells are not enriched for colony forming ability**

Although the basic phenotypic cellular hierarchy in the mammary gland has been established, much work remains to elucidate all of the progenitor cell populations. *In vitro* culture systems have been used extensively in the mammary field to assess both

differentiation and transformation potential. The NIH3T3 co-culture assay has been used by other groups to assess the ability of mammary cell populations to differentiate into myoepithelial and luminal epithelial cells.

Since this assay is a method to describe both the colony forming and differentiation potential of mammary cells, we began by validating the type of colonies produced by morphology. We stained the colonies formed in the co-culture assay with Wright-Giemsa after they had grown for 7 days in culture. We observed myoepithelial colonies contained cells that had more fibroblast-like phenotype with stellate morphology touching the bottom layer of irradiated NIH3T3 fibroblasts. Some cells retained their thin, spindle like appearance as observed *in vivo*, but most had a hybrid mesenchymal-epithelial morphology. Although these cells had mesenchymal characteristics, they were easily distinguishable from control NIH3T3 cells (Figure 2.4A, upper middle panel (myoepithelial), and upper left panel (3T3 alone)). Mature luminal epithelial colonies had compact morphology with defined boundaries (Figure 2.4A, upper right panel, luminal). These colonies had typical cultured epithelial growth upon closer examination (Figure 2.4A, lower right panel, luminal). Both myoepithelial and luminal epithelial colonies grew to similar sizes (Figure 2.4A, upper middle panel versus upper right panel). Our morphological analysis showed that the co-culture system was able to support growth of both luminal and myoepithelial colonies, which were easily distinguishable from each other.

To investigate the role of Thy-1 mammary cells, we tested the ability of Lineage$^-$, MRU, MaCFC, MYO and EPI cells to form colonies in the co-culture assay (Figure 2.4B). Lineage$^-$ cells formed an average of 90 colonies per 3000 cells plated, giving an

average frequency of 1 colony forming cell per 33 cells plated. Consistent with other reports[30,31], we found that both MRU and MYO populations were not adept at colony formation (Figure 2.2A). In contrast, MaCFCs were excellent at making colonies in this assay, yielding about 20 colonies for every 100 cells plated. These cells' ability to form colonies *in vitro* has been shown from their initial description[2]. EPIs were also adept at making colonies, but only half as good at MaCFCs, with an average of 52 colonies per 500 cells. All colony formation was dose dependent on the number of cells plated. Since neither MRU nor MYO cells produce many colonies when isolated and plated and were myoepithelial by morphology, we reason a majority of the colony formation in Lineage$^-$ cells comes from MaCFC and EPI cells.

Since we observed differential Thy-1 expression in MRU cells, we went on to test if Thy-1$^+$MRU or Thy-1$^-$MRU could enrich for colony forming cells as an indication of differentiation potential and compared the results to Lineage$^-$ and MaCFC colony formation. Neither Thy-1$^+$MRU nor Thy-1$^-$MRU populations had enrichment in colony formation (Figure 2.4C) compared to Lineage$^-$ or MaCFCs, suggesting Thy-1 does not enrich for MaCFC or EPI cells with the same phenotype as MRUs. It is likely that MYO cells are also isolated in the MRU gate, since these two populations share many characteristics. These results are also consistent with our data that Thy-1 is expressed in the myoepithelial compartment and previous reports that describe the mRNA expression of MRU to be similar to MYO cells[2,30].

**Thy-1$^+$CD24$^{med}$CD49f$^{hi}$ cells produce normal ductal epithelium**

Our *in vitro* results demonstrated Thy-1 did not isolate a colony forming sub-population, which was a characteristic of cells that had a luminal phenotype. Mammary stem cells express myoepithelial markers, and myoepithelial cells have poor colony forming potential *in vitro*. Recently, a number of reports have shown that Thy-1 enriches for stem cells in a variety of somatic tissues including blood, brain, liver, lung and general mesenchyme. We hypothesized Thy-1 could isolate the basal cell population enriched for functional mammary stem cells. To test for ductal regeneration potential, we performed limiting dilution transplants *in vivo*. The ability of prospectively isolated cell populations to produce ductal outgrowths in the cleared fat pads of recipient mice has long been used to successfully discriminate between mammary cell populations.

Due to the lack of consensus tissue processing method in the mammary field, different groups report variable mammary stem cell frequencies based on similar phenotype[2,3,30,32]. This has become a major hurdle for the mammary stem cell field. To address this issue, we attempted a number of different digestion protocols and engraftment strategies (data not shown). We found a hybrid digestion protocol based on reports from Sleeman, Shackleton and Stingl worked the best in our hands. To show our protocol was consistent with results from other groups, we compared the morphology of ducts from 6-week old wild type mice to transplanted 500 $CD24^{med}CD49f^{hi}$ (MRU) cells analyzed 10-weeks post-transplant (Figure 2.5A and B, respectively). Upon close inspection, we observed typical ductal structures and terminal lobules from our transplanted cells (Figure 2.5C), suggesting our technique did not adversely affect the ability of MRUs to develop in response to endogenous growth signals.

When unlabeled donor mammary cells are transplanted into weaning age mice, sometimes it may be difficult to distinguish ducts that have grown from incomplete clearing of the endogenous epithelium versus transplanted outgrowth epithelium. A key difference is endogenous ducts grow directionally first from the forth nipple area and then invade into the fat pad. Transplanted epithelium usually grows from a distant site deep within the fat pad. Thus, the site of injection may be used to distinguish the donor epithelium. Analyzing the site of injection also allowed us to see if donor epithelium was spatially and directionally regulated. We found transplanted epithelium typically grew in multiple directions from the site of injection (Figure 2.5D), indicating the entire mammary fat pad is able to support ductal growth, regardless of the site of transplantation. The initial site of transplantation was evident in all transplants analyzed.

To remove ambiguity from our transplants, all transplants of 800 cells and below were performed with pCx-GFP mammary cells, which ubiquitously expressed GFP in all ductal cells. We show (Figure 2.5E,F) the ductal outgrowths from a 25K Lineage$^-$ and a 30 cell Thy-1$^+$CD24$^{med}$CD49f$^{hi}$GFP$^+$ transplant as an example that the donor epithelium grew similarly to wild type ducts (Figure 2.5A), and produced normal ductal and TEB structures. We predict other cell types present in the Lineage$^-$ population may contribute to either engraftment or growth of ductal epithelium to fill the entire fat pad. Shackleton et al. co-injected supporting cells along with single stem cells but found no synergistic effect. This does not rule out that other populations such as preadipocytes or periductal cells could not provide this advantage.

**Thy-1$^+$CD24$^{med}$CD49f$^{hi}$ cells are enriched for mammary stem cells**

After proving our transplantation system, we performed an extensive limiting dilution *in vivo* transplantation study to determine the frequency of stem cells in the unsorted (Bulk), Lineage$^-$, Thy-1$^+$CD24$^{med}$CD49f$^{hi}$ and Thy-1$^-$CD24$^{med}$CD49f$^{hi}$ populations (Table 1.1). All transplants were performed with cells from donor 6-10 weeks old female mice (either wild type C57BL/6 or pCx-GFP transgenic reporter mice) transplanted into 21-28 day old female C57BL/6 recipient mice. For FACS based isolation of cell populations, we used the purity of the sorted cells to determines the quality of the results. We double sorted all populations before transplantation, and visualized GFP cells after sorting under a microscope to verify the cells were undamaged.

We found that unsorted mammary cells (Bulk) had a frequency of 1 stem cell in 46,876 cells (95% C.I. = 1 per 22,125 to 1 per 99,313 cells). Using flow cytometry to remove Lineage$^+$ expressing cells increased the stem cell frequency to 1 in 18,106 cells (95% C.I. = 1 per 11,994 to 1per 27,333 cells). The Lineage$^+$ cells constituted an average of 55.7% of Bulk cells, and removing these cells enriched 2.6-fold for stem cells over Bulk cells. We then isolated and injected CD24$^{med}$CD49f$^{hi}$ cells at limiting dilution to determine the frequency of mammary stem cells in our system. We found the frequency of stem cells defined by the CD24$^{med}$CD49f$^{hi}$ phenotype was 1 in 1,105 (95% C.I. = 1 per 534 to 1 per 2,285 cells). Isolating the CD24$^{med}$CD49f$^{hi}$ population gave a 16.4-fold increase of stem cells over Lineage$^-$ cells. We observed Thy-1 staining separated the Thy-1$^+$CD24$^{med}$CD49f$^{hi}$ population into Thy$^{hi}$ and Thy$^{low}$ fractions (Figure 2.6). Upon *in vivo* transplantation of these two populations, we found both had similar stem cell frequencies of 1 in 240 (95% C.I. = 1 per 114 to 1 per 504 cells) for Thy$^{hi}$ cells and 1 in 268 (95% C.I. of 1 per 192 to 1 per 374 cells) for Thy$^{low}$ cells. Therefore, we grouped these cells

together since there was no statistical difference in the frequency of stem cells between the two (Student's T-test, 2-tailed, p=0.7870). Strikingly, we found the Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ population had a frequency of 1 stem cell per 227 cells (95% C.I. = 1 per 167 to 1 per 310). This represented a 4.9-fold enrichment over the CD$24^{med}$CD$49f^{hi}$ (original MRU) population. The stem cell frequency of the Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ population was significantly enriched compared to the CD$24^{med}$CD$49f^{hi}$ (Student's T-test, 2-tailed, $p<0.0001$). In stark contrast, *in vivo* limiting dilution of the Thy-$1^{-}$ CD$24^{med}$CD$49f^{hi}$ population showed these cell were depleted of stem cells, with a frequency of 1 in 2,862 cells (95% C.I. = 1 per 1,255 to 1 per 6,526 cells). Although cells were not significantly different in stem cell frequency than the CD$24^{med}$CD$49f^{hi}$ population (Student's T-Test, 2-tailed, p=0.0896), they did have a 2.59-fold reduction of stem cells. Further analysis showed the Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ population was 12.6-fold enriched for stem cells over Thy-$1^{-}$CD$24^{med}$CD$49f^{hi}$ cells, and this difference was significant (Student's T-Test, 2-tailed, $p<0.0001$). The frequency of mammary stem cells in the transplanted populations may be found in Table 1.2.

**Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ ductal outgrowths have increased self-renewal**

We next investigated the ability of Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ and Thy-$1^{-}$ CD$24^{med}$CD$49f^{hi}$ cells to self-renew using secondary transplantation of portions of primary transplanted epithelium. Serial transplantation of donor epithelium represents the "gold standard" for measurement of cells' ability to self-renew, since only long lived populations will have to ability to regenerate the entire breast gland more than once. We found that small 1-3 mm chunks of primary ductal trees gave consistent secondary

outgrowths. Since we could not estimate the number of MRUs contained within each chunk engrafted, we calculated the frequency of secondary ductal growth. Our secondary transplants demonstrated Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ primary epithelium was enriched for self-renewal compared to Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ derived epithelium (Table 1.3). Lineage- primary ducts were easily able to self-renew, with 10 of 11 transplants successfully producing secondary epithelium. The glands that did form were morphologically normal, similar to our primary results. We transplanted pieces from both Thy-1$^{med}$CD24$^{med}$CD49f$^{hi}$ and Thy-1$^{hi}$CD24$^{med}$CD49f$^{hi}$ primary ducts to determine if there was a functional difference in the self-renewal of these two populations. Overall, Thy-1$^{med}$CD24$^{med}$CD49f$^{hi}$ primary epithelium had an engraftment efficiency of 68% and Thy-1$^{hi}$CD24$^{med}$CD49f$^{hi}$ had an efficiency of 75%. We did not find a significant difference in the self-renewal ability of primary epithelium from these two populations, confirming our previous conclusion that there is not a functional difference between Thy-1$^{hi}$ and Thy-1$^{med}$ expression on stem cells. The ability of primary ducts to self-renew was independent of the original numbers of cells transplanted, as epithelium from primary transplants of 600, 100 or 30 cells could self-renew. To test self-renewal of Thy-1$^{-}$ CD24$^{med}$CD49f$^{hi}$ cells, we transplanted pieces from 3K, 1K, and 300 cell primary outgrowths. The primary ductal outgrowths of Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells were often smaller and less developed than Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ primary trees (data not shown). When pieces were engrafted into recipient mice, we found only 1/9 secondary transplants were able to form secondary trees, indicating only an 11% engraftment efficiency. These data suggest Thy-1$^{-}$ CD24$^{med}$CD49f$^{hi}$ primary trees were five times more deficient in self-renewal ability as compared with Thy$^{+}$CD24$^{med}$CD49f$^{hi}$ primary ductal epithelium.

We dissociated the self-renewal transplants and stained for Thy-1, CD24 and CD49f to analyze the expression of these markers. All of our previously defined populations from freshly dissociated mammary glands were also present in secondary transplants (Figure 2.7A). Flow analysis revealed two major populations from secondary transplant, showing the core phenotypic cells that make up the mammary ductal system. Comparing Figure 2.7A with Figure 2.1B suggests the phenotype of stromal cells in the mammary gland. We also found Thy-1 heterogeneity in the $CD24^{med}CD49^{hi}$ cells (Figure 2.7B), which was an average of 22.4% positive cells. These results also prove $Thy\text{-}1^{+}CD24^{med}CD49^{hi}$ can reproduce all of the heterogeneity found in the mammary system, confirming the presence of stem cells.

When analyzed for morphology, secondary $Thy\text{-}1^{+}CD24^{med}CD49^{hi}$ ducts (Figure 2.7B) were normal and resembled $Lineage^{-}$ secondary transplants (Figure 2.7C). It should be noted that the only chunk that produced secondary trees were derived from a 1K cell $Thy\text{-}1^{-}CD24^{med}CD49^{hi}$ cells primary tree. Since flow cytometry is prone to small amounts of contaminating cells even when populations are double sorted, contamination with $Thy\text{-}1^{+}$MRU cells in the primary transplant may have produced a false positive tree. However, this growth of this tree was retarded, an unable to grow beyond a few cell divisions (Figure 2.7E). All results were all verified by Carmine Alum staining to assess if GFP expression was somehow silenced in these transplants (data not shown).

## DISCUSSION

The mammary stem cell phenotype based on expression of CD24 and CD49f has been described by a number of groups. Although phenotypic isolation of these cells has

led to their enrichment, further work remains to purify the population, allowing detailed transcriptional and protein analysis to discover stem cell specific regulatory pathways. Elucidating these mechanisms may prove valuable insights into oncogenesis[33]. Therefore, we undertook a study with the hypothesis that Thy-1, a GPI-anchored membrane protein that has been useful in isolating stem cells in other systems, would also be a marker of mammary stem cells. Our group has also shown that Thy-1 is a marker of cancer stem cells in the *MMTV-Wnt1* breast tumor mouse model[34], suggesting this protein is expressed preferentially in early progenitor cells. In addition, Thy-1 was identified as a marker of human mammary stem cells[35], suggesting Thy-1 mammary expression may be evolutionarily conserved in mammals.

Our study is the first detailed report of Thy-1 expression in the murine mammary gland. Using flow cytometry, we show that Thy-1 is highly expressed in the MRU population and decreases as cells become differentiated into luminal epithelium and myoepithelium. It should be noted that successful Thy-1 detection relies on using bright fluorophores coupled to the antibodies, such as APC or PE-Cy7, since the Thy-$1^{low}$ population may easily be misinterpreted as a negative population if flow cytometry selection gates are incorrectly calculated. We found only a small subset of Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ cells has a high amount of Sca-1 on their cell surface. However, since we did not perform limiting dilution *in vivo* transplantation based on Sca-1, we cannot assess its enrichment potential for mammary stem cells. There have been recent reports suggesting CD133[31] and CD61[36,37]} as a luminal-restricted progenitor marker. Using these additional markers may help to enrich for mammary stem cells, either by positively selected for the stem cells or depleting the more differentiated progenitors. In

addition, Shackleton and colleagues have described CD29 as a mammary stem cell marker[3]. We found the $CD24^{med}CD49f^{hi}$ population was a subset of the $CD24^{+}CD29^{hi}$ population, and the reported frequency of stem cells was higher in the $CD24^{med}CD49f^{hi}$ population. In addition, the two antibodies we used for CD29 (HMG1.1 and Ha2-5) were dim, reducing the effective separation of these cells from other channels in FACS. Since there is considerable overlap between CD29 and CD49f expression in mammary epithelial cells, we did not we lost significant information by excluding this marker.

We show that EPI cells are excellent at colony formation in co-culture assay, but not as efficient as MaCFCs. Although our data show Thy-1 is differentially expressed in mammary stem and progenitor cells, Thy-1 did not isolate a fraction of the $CD24^{med}CD49f^{hi}$ population that was similar to further differentiated progenitor cells with robust ability to form colonies *in vitro*, suggesting these cells were myoepithelial cells or shared characteristics of basal cells. Recently, there has been description of a new *in vitro* assay to assess self-renewal of mammary cells. The mammosphere assay is essentially an adaptation of the neurosphere assay, which is based on evidence that suggests neural stem cells are able to grow in suspension (i.e. purposefully not allowed to attach to a culture surface) in a clonal fashion, retaining their self-renewal and differentiation potential[38]. This assay has been claimed as an *in vitro* method to estimate the number of self-renewing cells in a population based on the ability to form secondary spheres from dissociated primary ones. We chose not to perform mammosphere assays on our isolated cells because *in vivo* transplants provided stronger evidence of the enrichment for mammary stem cells in a phenotypic population. In addition, the mammosphere assay has not been properly investigated to warrant such bold

claims in the mammary system. For example, there is no agreement on how many cells in a culture constitute clonal density so that each mammosphere is derived from a single cell. Therefore, questions about aggregation and clonality plague this system. In addition, a recent study showed that individual mammospheres have a 15% engraftment efficiency *in vivo*, in stark contrast to the primary single cell engraftment efficient from which they arise[39]. Culturing the cells may have adverse effects on maintenance of self-renewal. This point leads to the lack of validation of how many cells are actually self-renewing in each sphere formed. Two studies have shown that the classical neurosphere assay overestimates the frequency of stem cell number by 30-50 fold[40,41]. Given the ambiguity of the cells that actually form spheres, aggregation issues, and inhibited *in vivo* engraftment after culture, this assay has failed as an accurate assessment of either mammary stem cell self-renewal or differentiation. A simple, yet conspicuously overlooked experiment to prove the validity of the system is limiting dilution *in vivo* transplantation of dissociated cells from primary mammospheres versus secondary and tertiary spheres, showing at least an increase or maintenance or stem cell frequency. Without such data, reports of murine mammary gland cell utilizing the mammosphere assay must be subject to increased scrutiny. The human mammosphere assay has fared better in general scientific acceptance, yet still suffers from the same drawbacks as the murine mammosphere system. However, results have shown that individual spheres are able to produce human ductal structures *in vivo*[42,43]. In addition, experiments using combinations of marked human mammary cells plated into the mammosphere assay at extremely low density produce clonal spheres[44].

Our conclusion is that Thy-1 marks a basal cell population and was supported by histological analysis of Thy-1 expression, which was localized near the basal lamina in mature ducts. Although Thy-1 did not isolate a MaCFC-like progenitor population, we went on to investigate if Thy-1 enriched mammary stem cells in limiting dilution transplantation. Our *in vivo* results demonstrated the Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ population was significantly enriched for mammary stem cells compared to their frequency in CD$24^{med}$CD$49f^{hi}$ (MRU) cells. We also found the Thy-$1^{-}$CD$24^{med}$CD$49f^{hi}$ had a severe depletion of mammary stem cells. We were unable to produce any in vivo outgrowths from the MaCFC population, even when 25K cells were transplanted (data not shown). We went on to test if Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ derived epithelium could self-renew using serial transplantation. Our data shows that Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ cells can both self-renew and reproduce the heterogeneity found in freshly dissociated mammary tissue. Ductal morphogenesis of secondary was comparable to wild type mammary glands. Thy-$1^{-}$CD$24^{med}$CD$49f^{hi}$ primary epithelium had significantly decreased ability to produce secondary outgrowths.

The initial report by Stingl and colleagues reported a much higher frequency of stem cells per gland, and reported the MaCFC population also contained MRUs. We were unable to reproduce their stem cell frequencies, even identical reagents and rigorous attention to their described protocol. We did find the phenotypic CD$24^{med}$CD$49f^{hi}$ population they described was enriched for stem cells versus all other populations tested (data not shown). Also, in Stingl et al. the transplant model system involved mating the recipient mice two weeks after initial surgery, which we did not perform. We chose not to mate our mice post-transplant since the hormonal profile of the

mouse changes dramatically, stimulating massive expansion of the mammary system and creating a growth advantage for specifically for alveolar progenitor cells, the source of a potential parity-induced second stem cell population that may provide false-positive results. Since these mechanisms are poorly understood, we could not anticipate how pregnancy would affect either niche spaces or stem cell expansion.

Collectively, our data demonstrates Thy-1 is a mammary stem cell marker. In the blood system, the major difference between hematopoietic stem cells (HSCs) and multipotent progenitor cells (MPPs) is self-renewal. MPPs have the same differentiation potential as HSCs, but they are only able to transiently reconstitute the blood system of transplanted mice. Our data shows the Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells fit the criteria for multipotent progenitor parameters. Although they have the ability to engraft in mice (although much less than the Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ cells) and produce primary ductal epithelium, they have little to no self-renewal activity. Therefore, we propose that the Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells are mammary multipotent progenitor cells (MMPP), a population analogous to MPP cells in the blood system. We find these cells have a closer relationship to the mammary stem cells than the MaCFC population since they share basal/myoepithelial characteristics with the stem cells and are able to engraft *in vivo*.

Although good markers such as Thy-1, CD24, CD29 and CD49f have been used to describe MRUs, further work is warranted to purify the stem cell population. In the rat mammary system, peanut agglutinin (PNA) has been successfully used to isolate clonogenic cells that have an enriched ability to give rise to alveolar units in hyperprolactinemic recipients, suggesting rat multipotent stem cells are PNA+. This protein has yet to be explored in the mouse mammary system, although flow cytometry

analyses shows this marker may be similar in expression to CD24[22]. In addition, peanut lectin has been used as a marker of luminal epithelial mammary cells *in vivo[45]*. These are two examples of potential mammary stem cell markers that have yet to be investigated. Future work in the field will allow single cell clonal analysis of mammary populations, warranting studies on how different mammary populations may be corrupted in oncogenesis.

**TABLES**

**Table 2.1: Limiting dilution transplants of mammary cells into the clear fat pads of syngeneic recipient mice.**

| | Bulk | Lin$^-$ | CD24$^{med}$CD249f$^{hi}$ | Thy-1$^+$ CD24$^{med}$CD49f$^{hi}$ | Thy-1$^-$ CD24$^{med}$CD49$^{hi}$ |
|---|---|---|---|---|---|
| 250K | 1/1 | | | | |
| 133K | 2/2 | | | | |
| 100K | 3/4 | | | | |
| 50K | | 3/3 | | | |
| 25K | 4/8 | 11/18 | | | |
| 20K | | 3/4 | 2/2 | | |
| 10K | | 9/15 | 2/2 | | |
| 5K | | 1/4 | | | |
| 3K | | | 1/3 | 2/2 | 1/2 |
| 2K | | | 1/2 | | |
| 1K | | | 4/4 | 3/7 | 1/3 |
| 800 Cells | | | 1/1 | | |
| 600 Cells | | | | 1/2 | 0/2 |
| 500 Cells | | | 1/4 | | 0/1 |
| 400 Cells | | | | 2/4 | 0/4 |
| 300 Cells | | | 2/3 | 1/6 | 1/3 |
| 250 Cells | | | | 2/2 | 0/3 |
| 220 Cells | | | | | 0/1 |
| 200 Cells | | | | 5/11 | 0/5 |
| 100 Cells | | | | 25/44 | 2/30 |
| 75 Cells | | | | | 0/1 |
| 70 Cells | | | 0/1 | | |
| 50 Cells | | | | 13/31 | 1/24 |
| 30 Cells | | | | 12/15 | 0/9 |
| 24 Cells | | | | 0/1 | |
| 21 Cells | | | | | 0/1 |
| 17 Cells | | | | 0/1 | |

**Table 2.2: Frequency of mammary stem cells in isolated populations.**

| **Population** | **Frequency** |
|---|---|
| Bulk | 0.0000213 |
| Lineage$^{-}$ | 0.0000552 |
| CD24$^{med}$CD49f$^{hi}$ | 0.000905 |
| Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ | 0.004405 |
| Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ | 0.000349 |

**Table 1.3: Secondary transplants with pieces of primary Thy-1$^{+}$CD24$^{med}$CD49$^{hi}$ ductal outgrowth epithelium.**

| Original Transplanted Population | Engrafted/Transplanted | Efficiency |
|---|---|---|
| Lineage$^{-}$ 25K | 10/11 | 91% |
| Thy-1$^{med}$CD24$^{med}$CD49f$^{hi}$ 600 cells | 5/6 | 83% |
| Thy-1$^{med}$CD24$^{med}$CD49f$^{hi}$ 100 cells | 6/10 | 60% |
| Thy-1$^{med}$CD24$^{med}$CD49f$^{hi}$ 30 cells | 6/9 | 67% |
| Thy-1$^{hi}$CD24$^{med}$CD49f$^{hi}$ 100 cells | 6/8 | 75% |
| Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells | 1/9 | 11% |

## FIGURES

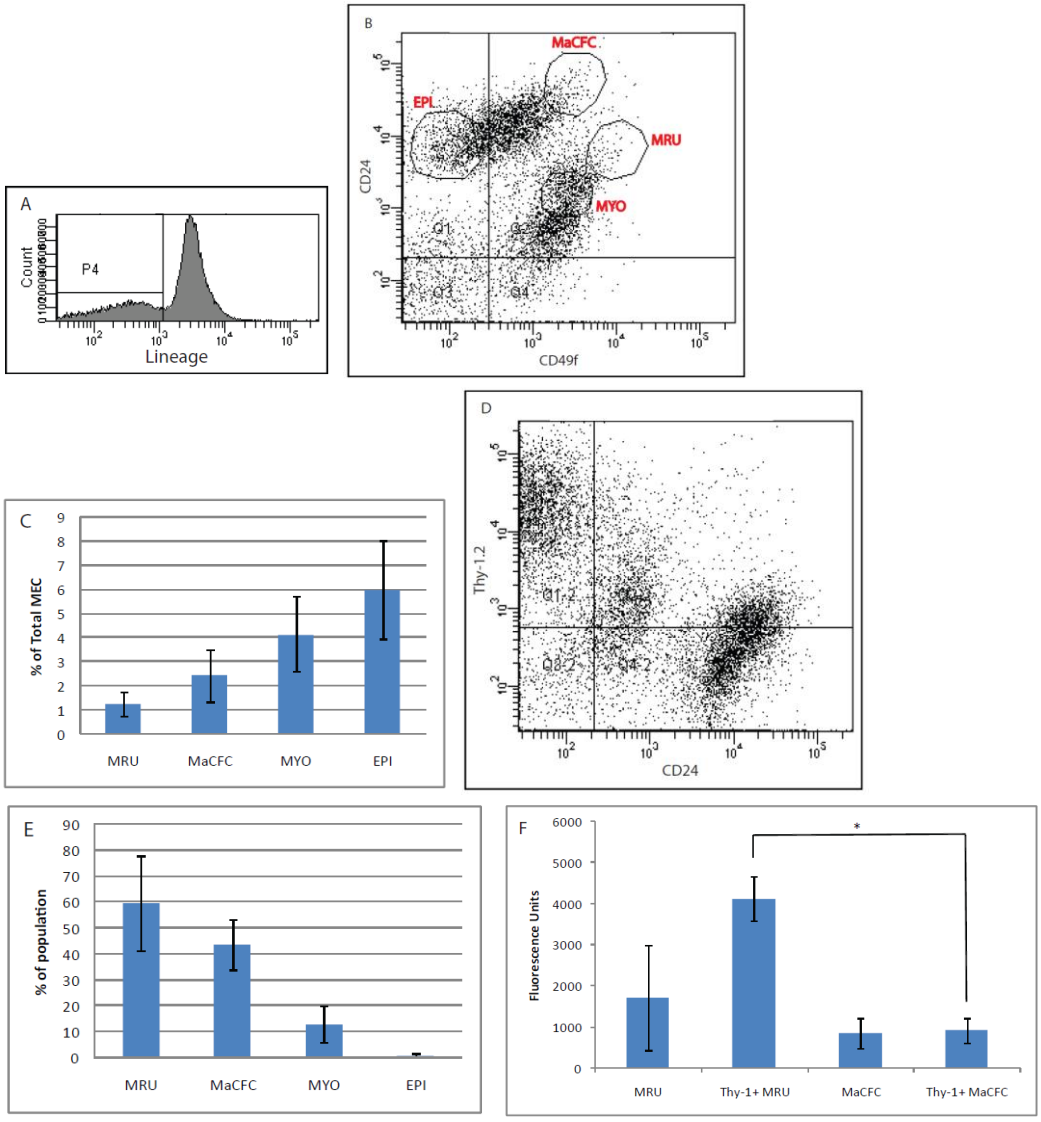



**Figure 2.1: Expression of Thy-1 on mammary epithelial cells.**

Single cell suspensions of mammary cells were stained with antibodies against Thy-1 and CD24. A, Representative histogram plot showing gating for removal of cells positive for lineage antibodies. B,D Representative flow cytometry dot plots of mammary cells gating out contaminating Lineage$^{+}$ cells. C, Percentages of mammary cell populations as defined by A. E, Thy-1 positive percentage of individual populations. F, average mean fluorescence intensity of Thy-1$^{+}$ and Thy-1$^{-}$ MRU and MaCFC populations as an indirect measurement the amount of protein found on the surface of the cells. Results from D, E and F are derived from 8 independent experiments.

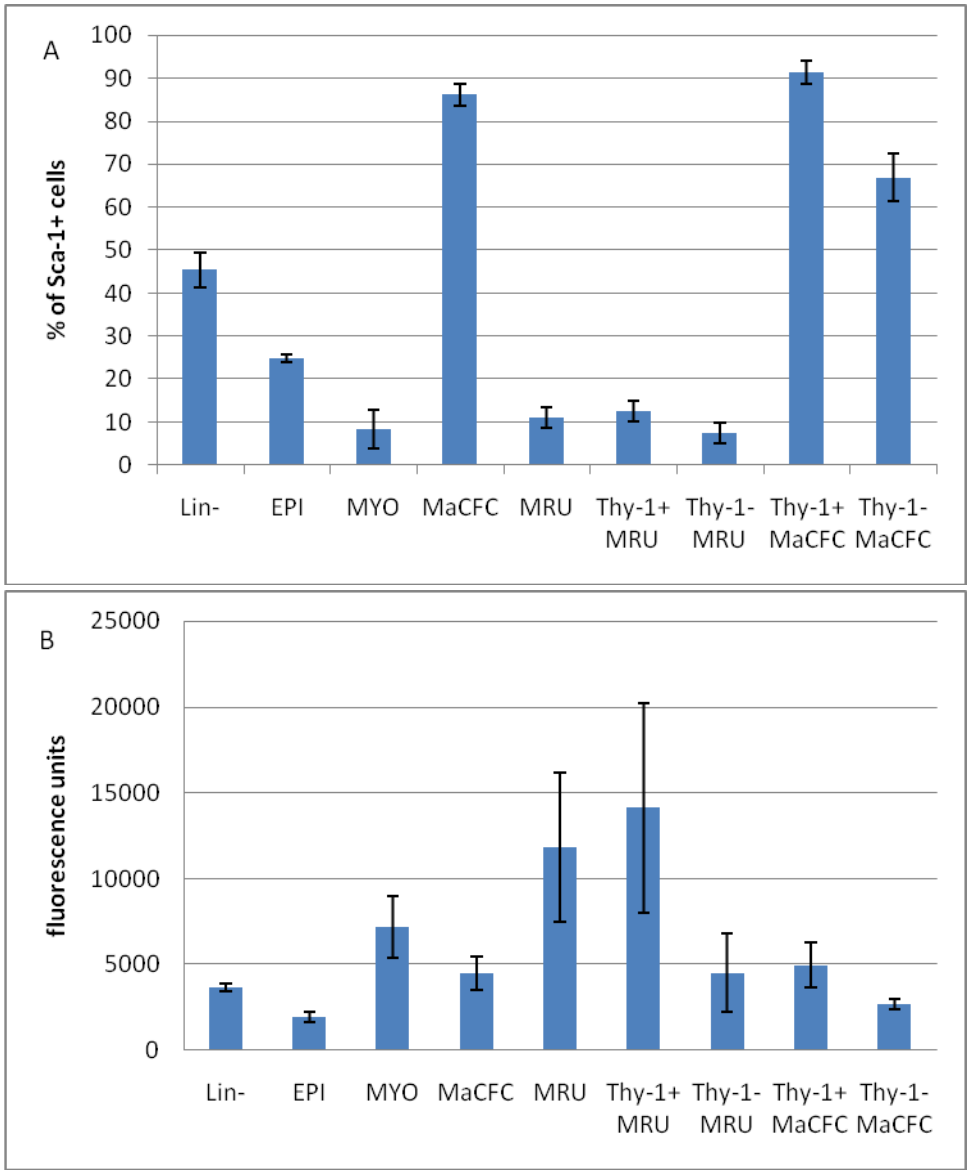



**Figure 2.2: Sca-1 expression and intensity on mammary epithelial populations.**

A, percentage of Sca-1 positive cells in each population. B, mean fluorescence intensity of Sca-1 positivity on individual populations as an indirect measurement of the amount of protein found on the surface of cells.

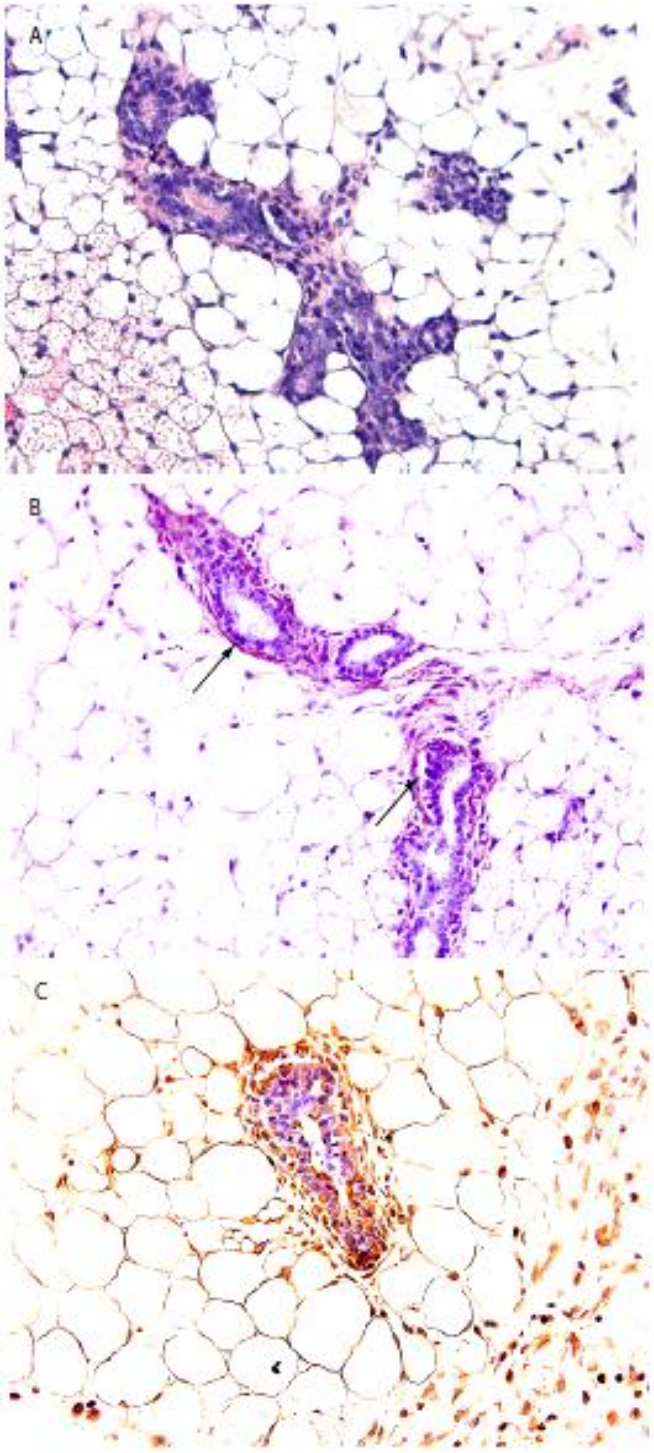



**Figure 2.3: Histology of mammary ducts in virgin mice.**

A, H&E staining. B, Thy-1 staining. C, CD24 staining. Thy-1 localizes to cells near the basal lamina, and CD24 can be found in periductal, myoepithelial and luminal cells. All images were taken at 25X magnification magnification.

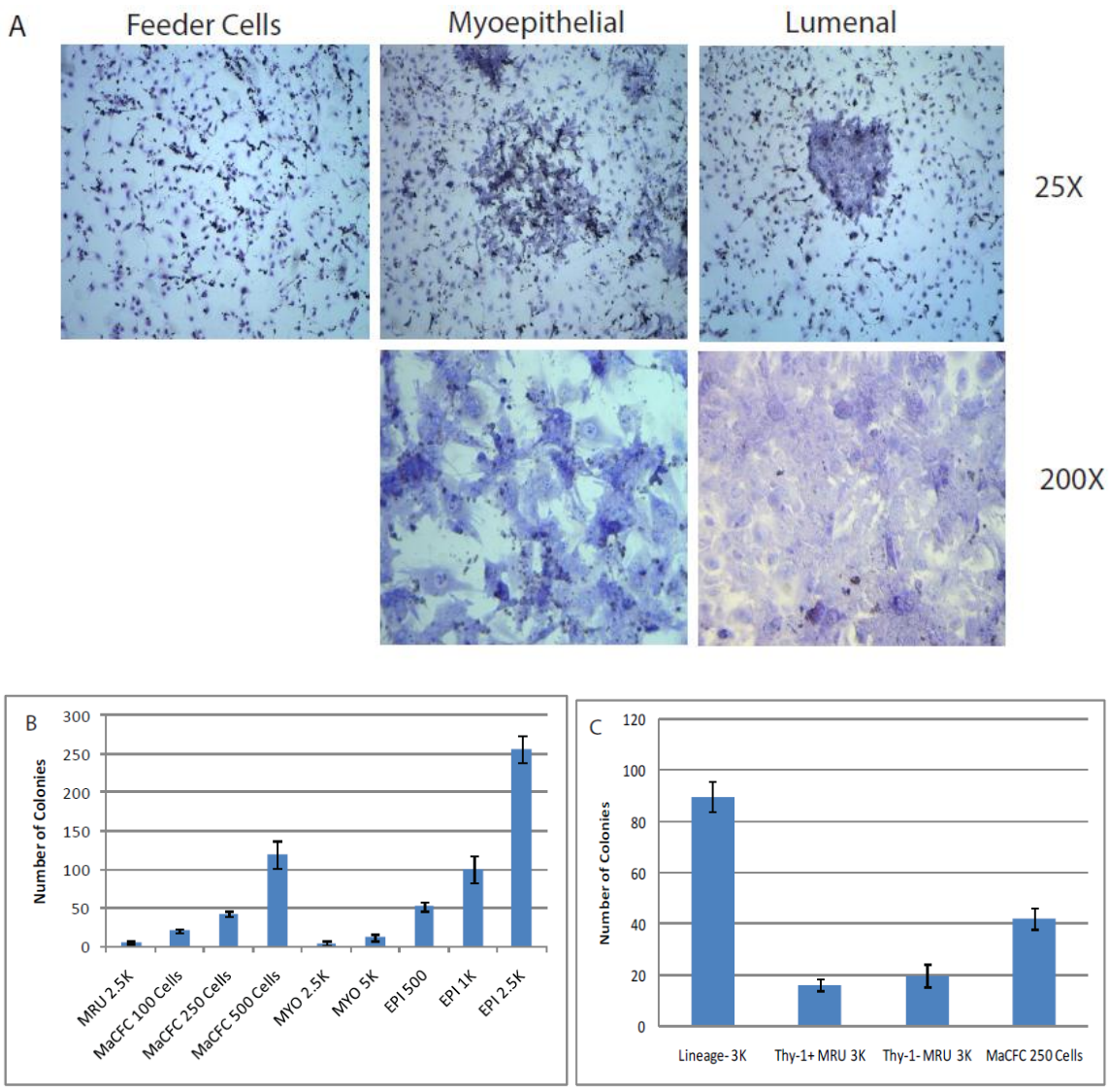



**Figure 2.4: Thy-1 does not enrich the stem cell phenotype for colony forming cells.**

A. Double sorted cell populations were plated on top of irradiated NIH3T3 fibroblast feeder layer cells. The resulting colonies were stained with Wright-Giemsa and assessed for morphology. Myoepithelial and luminal cell colonies are shown. B, Number of colonies formed from isolated mammary populations. C, colony formation of cells isolated on MRU phenotype ($CD24^{med}CD49f^{hi}$) and further separated based on Thy-1 expression. Results are based on 3 independent experiments.

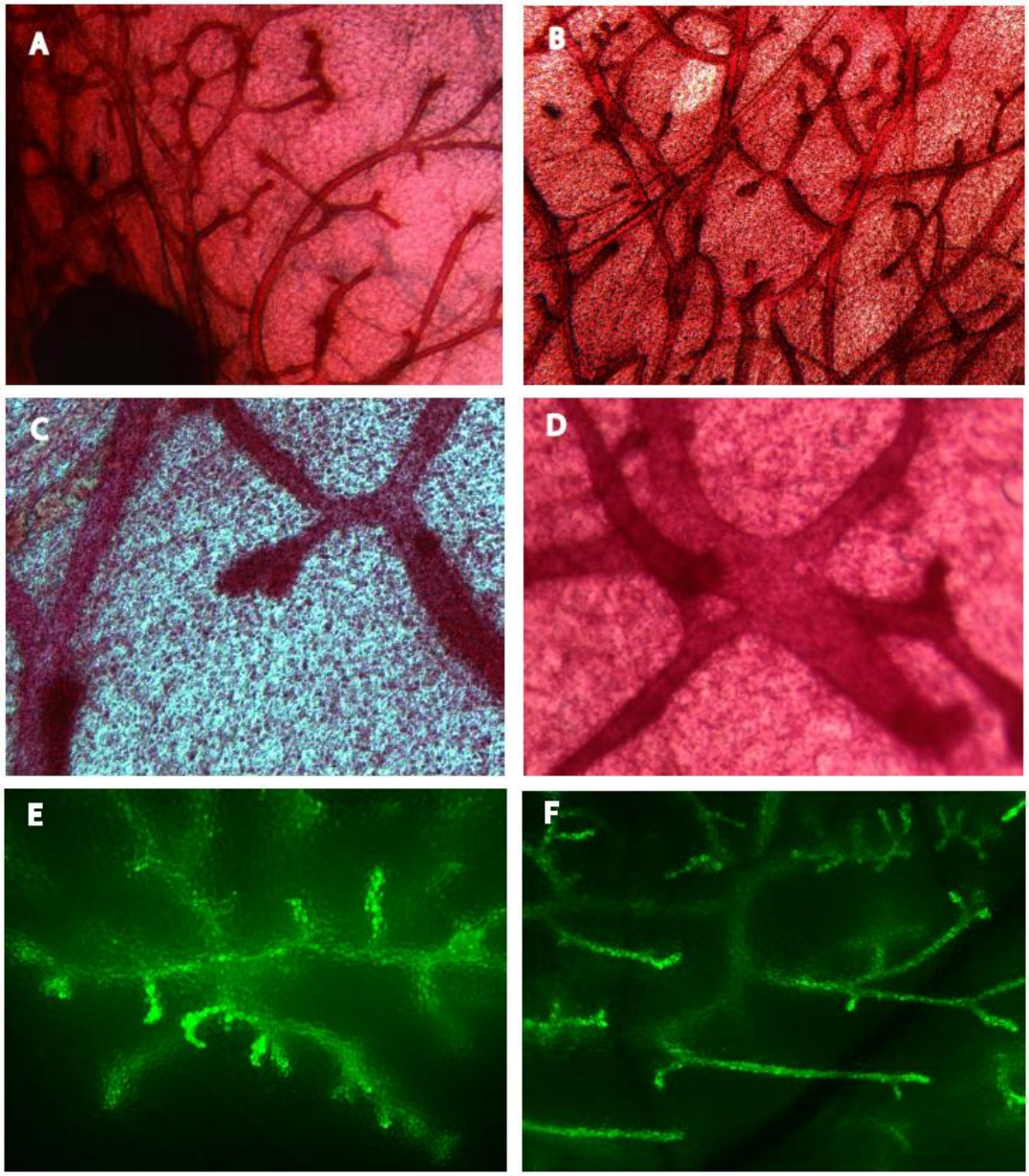



**Figure 2.5: Transplanted mammary cells have normal ductal outgrowth morphology.**

A, Whole mount of mammary gland from 6-week old C57Bl/6 female mouse. 25X magnification. B, Whole mount of 500 $CD24^{med}CD49f^{hi}$ transplanted cells after 10 weeks in recipient mouse. 25X magnification. C, Carmine stained whole mount from transplanted cells. 100X magnification. D, Original site of transplanted cells in recipient cleared fat pad. E, Ductal outgrowth from 30 $Thy\text{-}1^{+}CD24^{med}CD49f^{hi}GFP^{+}$ transplanted cells. F, Ductal outgrowth from 25K $Lineage^{-}GFP^{+}$ donor cells.

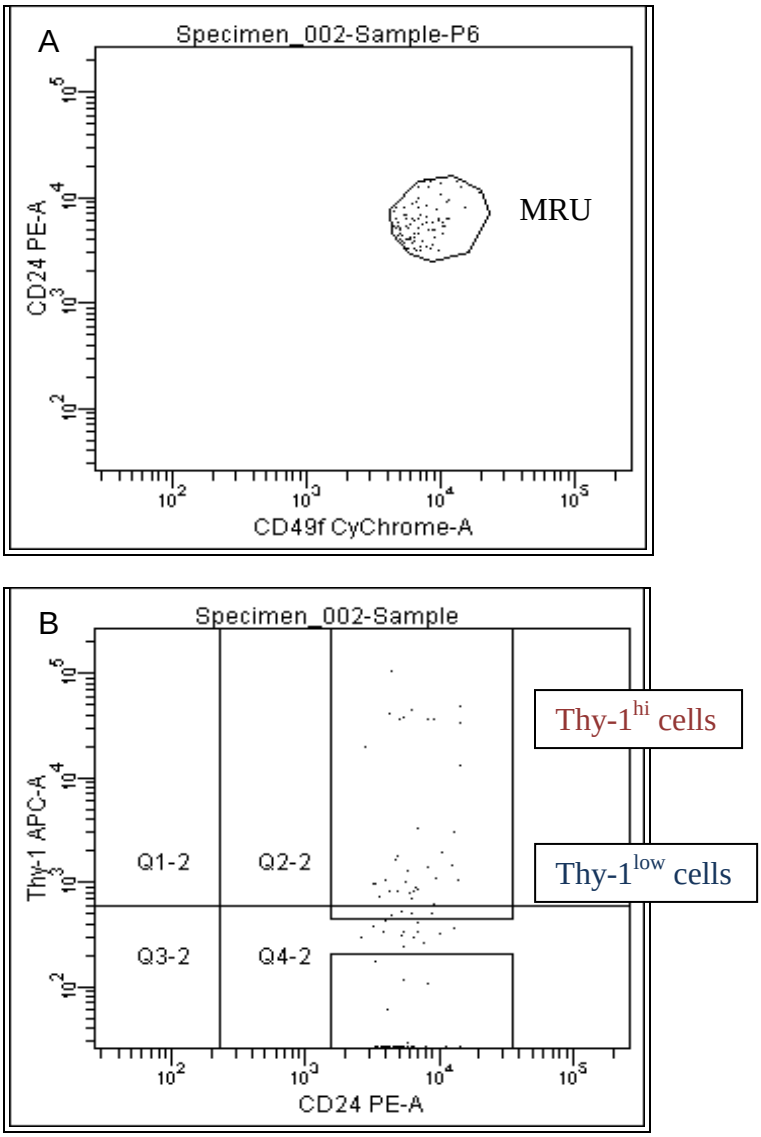



**Figure 2.6: Thy-1$^{hi}$ and Thy$^{low}$ cells in the Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ population.**

A. Gate for CD24$^{med}$CD49f$^{hi}$ population (MRU) that is enriched for mammary stem cells.
B. Thy-1 and CD24 expression based on the MRU gate.

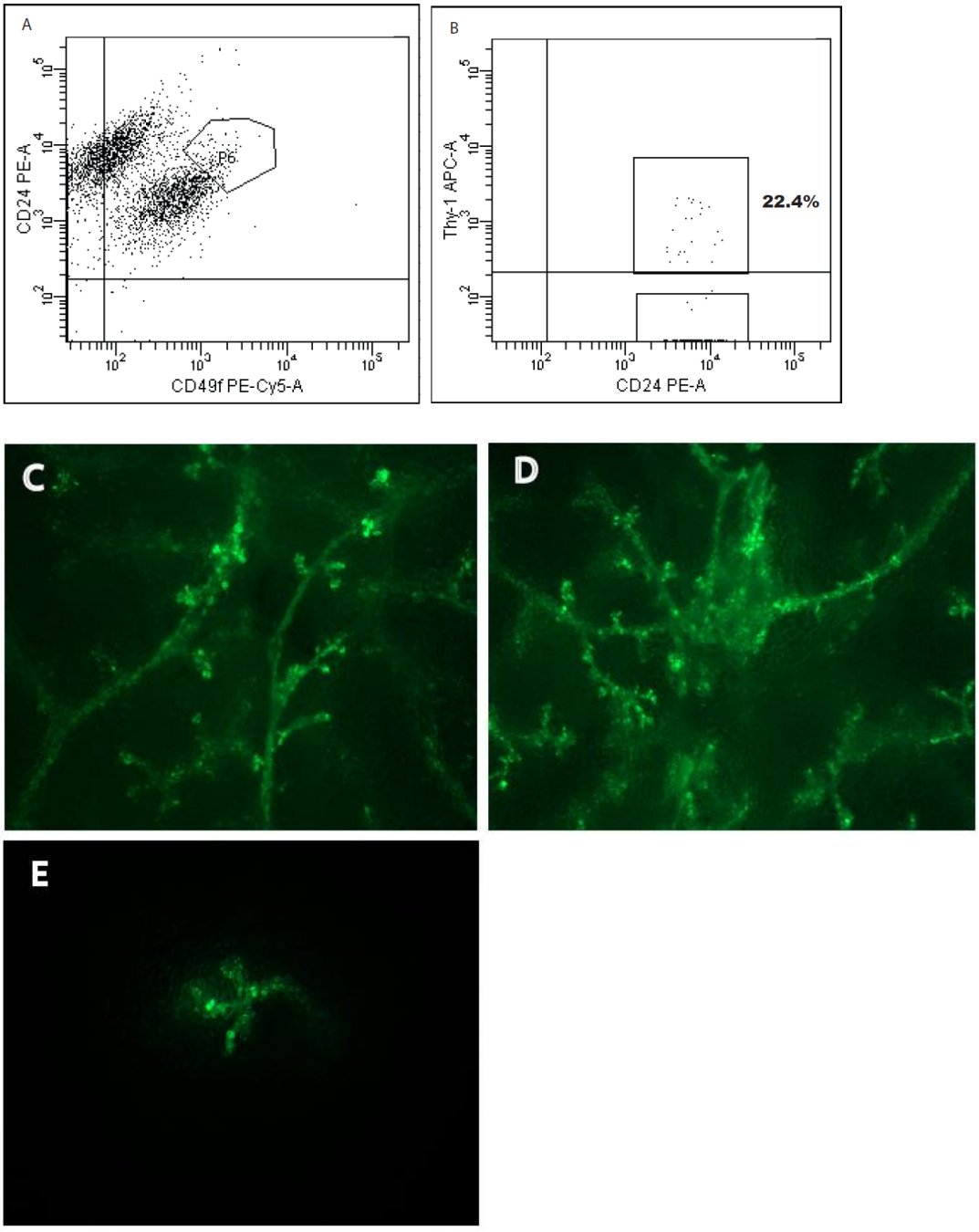



**Figure 2.7: Self-renewal of primary mammary ductal epithelium.**

A,B flow cytometry dot plots showing expression of CD24, CD49f and Thy-1 on dissociated cells obtained from Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ secondary transplants. C, secondary transplant of 25K Lineage$^{-}$ cells. 25X magnification. D, secondary transplant of 30 cell Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ primary epithelium. 25X magnification. E, only secondary transplant from 1K Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ primary epithelium. 25X magnification.

**BIBLIOGRAPHY**


1. Park PC, Selvarajah S, Bayani J, Zielenska M, Squire JA: **Stem cell enrichment approaches**. *Semin Cancer Biol* 2007, **17**:257-264.
2. Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, Li HI, Eaves CJ: **Purification and unique properties of mammary epithelial stem cells**. *Nature* 2006, **439**:993-997.
3. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, Wu L, Lindeman GJ, Visvader JE: **Generation of a functional mammary gland from a single stem cell**. *Nature* 2006, **439**:84-88.
4. Campbell DG, Gagnon J, Reid KB, Williams AF: **Rat brain Thy-1 glycoprotein. The amino acid sequence, disulphide bonds and an unusual hydrophobic region**. *Biochem J* 1981, **195**:15-30.
5. Reif AE, Allen JM: **The Akr Thymic Antigen and Its Distribution in Leukemias and Nervous Tissues**. *J Exp Med* 1964, **120**:413-433.
6. Goldschneider I, Gordon LK, Morris RJ: **Demonstration of Thy-1 antigen on pluripotent hemopoietic stem cells in the rat**. *J Exp Med* 1978, **148**:1351-1366.
7. Boswell HS, Wade PM, Jr., Quesenberry PJ: **Thy-1 antigen expression by murine high-proliferative capacity hematopoietic progenitor cells. I. Relation between sensitivity to depletion by Thy-1 antibody and stem cell generation potential**. *J Immunol* 1984, **133**:2940-2949.
8. Miller BA, Lipton JM, Linch DC, Burakoff SJ, Nathan DG: **Thy-1 is a differentiation antigen that characterizes immature murine erythroid and myeloid hematopoietic progenitors**. *J Cell Physiol* 1985, **123**:25-32.
9. Spangrude GJ, Brooks DM: **Phenotypic analysis of mouse hematopoietic stem cells shows a Thy-1-negative subset**. *Blood* 1992, **80**:1957-1964.
10. Uchida N, Weissman IL: **Searching for hematopoietic stem cells: evidence that Thy-1.1lo Lin- Sca-1+ cells are the only stem cells in C57BL/Ka-Thy-1.1 bone marrow**. *J Exp Med* 1992, **175**:175-184.
11. Craig W, Kay R, Cutler RL, Lansdorp PM: **Expression of Thy-1 on human hematopoietic progenitor cells**. *J Exp Med* 1993, **177**:1331-1342.
12. Peault B, Weissman IL, Buckle AM, Tsukamoto A, Baum C: **Thy-1-expressing CD34+ human cells express multiple hematopoietic potentialities in vitro and in SCID-hu mice**. *Nouv Rev Fr Hematol* 1993, **35**:91-93.
13. Oertel M, Menthena A, Chen YQ, Shafritz DA: **Comparison of hepatic properties and transplantation of Thy-1(+) and Thy-1(-) cells isolated from embryonic day 14 rat fetal liver**. *Hepatology* 2007, **46**:1236-1245.
14. Masson NM, Currie IS, Terrace JD, Garden OJ, Parks RW, Ross JA: **Hepatic progenitor cells in human fetal liver express the oval cell marker Thy-1**. *Am J Physiol Gastrointest Liver Physiol* 2006, **291**:G45-54.
15. Eslaminejad MB, Nadri S, Hosseini RH: **Expression of Thy 1.2 surface antigen increases significantly during the murine mesenchymal stem cells cultivation period**. *Dev Growth Differ* 2007, **49**:351-364.

16. Barlow JZ, Huntley GW: **Developmentally regulated expression of Thy-1 in structures of the mouse sensory-motor system**. *J Comp Neurol* 2000, **421**:215-233.
17. Stohl W, Gonatas NK: **Distribution of the thy-1 antigen in cellular and subcellular fractions of adult mouse brain**. *J Immunol* 1977, **119**:422-427.
18. Hilgers J, van Blitterswijk WJ, Bont WS, Theuns GJ, Nusse R, Haverman J, Emmelot P: **Distribution and antibody-induced redistribution of a mammary tumor virus-induced and a normal antigen on the surface of mouse leukemia cells**. *J Natl Cancer Inst* 1975, **54**:1335-1342.
19. Hilgers J, Haverman J, Nusse R, van Blitterswijk WJ, Cleton FJ, Hageman PC, van Nie P, Calafat J: **Immunologic, virologic, and genetic aspects of mammary tumor virus-induced cell-surface antigens: presence of these antigens and the Thy 1.2 antigen on murine mammary gland and tumor cells**. *J Natl Cancer Inst* 1975, **54**:1323-1333.
20. Dulbecco R, Elkington J: **Induction of growth in resting fibroblastic cell cultures by Ca++**. *Proc Natl Acad Sci U S A* 1975, **72**:1584-1588.
21. Rudland PS, Weil S, Hunter AR: **Changes in RNA metabolism and accumulation of presumptive messenger RNA during transition from the growing to the quiescent state of cultured mouse fibroblasts**. *J Mol Biol* 1975, **96**:745-766.
22. Kim ND, Oberley TD, Clifton KH: **Primary culture of flow cytometry-sorted rat mammary epithelial cell (RMEC) subpopulations in a reconstituted basement membrane, Matrigel**. *Exp Cell Res* 1993, **209**:6-20.
23. Dezso K, Jelnes P, Laszlo V, Baghy K, Bodor C, Paku S, Tygstrup N, Bisgaard HC, Nagy P: **Thy-1 is expressed in hepatic myofibroblasts and not oval cells in stem cell-mediated liver regeneration**. *Am J Pathol* 2007, **171**:1529-1537.
24. Boulanger CA, Wagner KU, Smith GH: **Parity-induced mouse mammary epithelial cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression**. *Oncogene* 2005, **24**:552-560.
25. Barraclough R, Kimbell R, Rudland PS: **Differential control of mRNA levels for Thy-1 antigen and laminin in rat mammary epithelial and myoepithelial-like cells**. *J Cell Physiol* 1987, **131**:393-401.
26. Lennon VA, Unger M, Dulbecco R: **Thy-1: a differentiation marker of potential mammary myoepithelial cells in vitro**. *Proc Natl Acad Sci U S A* 1978, **75**:6093-6097.
27. Welm B, Behbod F, Goodell MA, Rosen JM: **Isolation and characterization of functional mammary gland stem cells**. *Cell Prolif* 2003, **36 Suppl 1**:17-32.
28. Welm BE, Tepera SB, Venezia T, Graubert TA, Rosen JM, Goodell MA: **Sca-1(pos) cells in the mouse mammary gland represent an enriched progenitor cell population**. *Dev Biol* 2002, **245**:42-56.
29. Klinowska TC, Alexander CM, Georges-Labouesse E, Van der Neut R, Kreidberg JA, Jones CJ, Sonnenberg A, Streuli CH: **Epithelial development and differentiation in the mammary gland is not dependent on alpha 3 or alpha 6 integrin subunits**. *Dev Biol* 2001, **233**:449-467.
30. Sleeman KE, Kendrick H, Ashworth A, Isacke CM, Smalley MJ: **CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells**. *Breast Cancer Res* 2006, **8**:R7.

31. Sleeman KE, Kendrick H, Robertson D, Isacke CM, Ashworth A, Smalley MJ: **Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland**. *J Cell Biol* 2007, **176**:19-26.
32. Pietersen AM, Evers B, Prasad AA, Tanger E, Cornelissen-Steijger P, Jonkers J, van Lohuizen M: **Bmi1 regulates stem cells and proliferation and differentiation of committed cells in mammary epithelium**. *Curr Biol* 2008, **18**:1094-1099.
33. Lobo NA, Shimono Y, Qian D, Clarke MF: **The biology of cancer stem cells**. *Annu Rev Cell Dev Biol* 2007, **23**:675-699.
34. Cho RW, Wang X, Diehn M, Shedden K, Chen GY, Sherlock G, Gurney A, Lewicki J, Clarke MF: **Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors**. *Stem Cells* 2008, **26**:364-371.
35. Eirew P, Stingl J, Raouf A, Turashvili G, Aparicio S, Emerman JT, Eaves CJ: **A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability**. *Nat Med* 2008, **14**:1384-1389.
36. Luo M, Fan H, Nagy T, Wei H, Wang C, Liu S, Wicha MS, Guan JL: **Mammary epithelial-specific ablation of the focal adhesion kinase suppresses mammary tumorigenesis by affecting mammary cancer stem/progenitor cells**. *Cancer Res* 2009, **69**:466-474.
37. Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, Hartley L, Robb L, Grosveld FG, van der Wees J, et al.: **Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation**. *Nat Cell Biol* 2007, **9**:201-209.
38. Dontu G, Abdallah WM, Foley JM, Jackson KW, Clarke MF, Kawamura MJ, Wicha MS: **In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells**. *Genes Dev* 2003, **17**:1253-1270.
39. Moraes RC, Zhang X, Harrington N, Fung JY, Wu MF, Hilsenbeck SG, Allred DC, Lewis MT: **Constitutive activation of smoothened (SMO) in mammary glands of transgenic mice leads to increased proliferation, altered differentiation and ductal dysplasia**. *Development* 2007, **134**:1231-1242.
40. Reynolds BA, Rietze RL: **Neural stem cells and neurospheres--re-evaluating the relationship**. *Nat Methods* 2005, **2**:333-336.
41. Louis SA, Rietze RL, Deleyrolle L, Wagey RE, Thomas TE, Eaves AC, Reynolds BA: **Enumeration of neural stem and progenitor cells in the neural colony-forming cell assay**. *Stem Cells* 2008, **26**:988-996.
42. Villadsen R, Fridriksdottir AJ, Ronnov-Jessen L, Gudjonsson T, Rank F, LaBarge MA, Bissell MJ, Petersen OW: **Evidence for a stem cell hierarchy in the adult human breast**. *J Cell Biol* 2007, **177**:87-101.
43. Liu S, Dontu G, Mantle ID, Patel S, Ahn NS, Jackson KW, Suri P, Wicha MS: **Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malignant human mammary stem cells**. *Cancer Res* 2006, **66**:6063-6071.
44. Dontu G, Jackson KW, McNicholas E, Kawamura MJ, Abdallah WM, Wicha MS: **Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells**. *Breast Cancer Res* 2004, **6**:R605-615.
45. Newman RA, Klein PJ, Rudland PS: **Binding of peanut lectin to breast epithelium, human carcinomas, and a cultured rat mammary stem cell: use of the lectin**

**as a marker of mammary differentiation**. *J Natl Cancer Inst* 1979, **63**:1339-1346.

# CHAPTER 3

# MICROARRAY ANALYSIS OF MURINE MAMMARY POPULATIONS

## SUMMARY

Our recent description of murine mammary populations that were highly enriched for stem cells (MaSC) based on the Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ phenotype and for multipotent progenitors (MMPP) based on the Thy-1$^{-}$CD24$^{med}$CD49$^{hi}$ phenotype, enabled us to profile their global gene expression. Cytokeratin cluster analysis showed that the MaSC and MMPP populations expressed a myoepithelial signature. Real-time PCR validation of microarray results led to the discovery that p21$^{cip1/waf1}$ was expressed specifically in MaSCs and Tbx3 in further differentiated progenitors. *Ex vivo* protein analysis of sorted cells also confirmed our microarray results and suggested the MaSCs and MMPPs express proteins of both luminal and myoepithelial lineages. We also reported the first description of murine Epithelial Specific Antigen (ESA), a long sought after marker that discriminates between epithelium and stroma. Immunofluorescence staining of mature ducts showed that Krt19 and Krt6 expression was confined to the luminal compartment. Gene Set Enrichment Analysis demonstrated MaSCs have different molecular functions activated as opposed to MYO cells. Using genes expressed

by MMPP, MaCFC, MYO and EPI cells as filters, we presented a gene signature for MaSC cells. Gene ontology analysis of this stem cell signature showed the MaSCs have activated ion homeostasis, adhesion, motility and ATP production pathways. These results provided the first insights into the behavior of mammary stem cells.

## INTRODUCTION

Many groups have recently reported major advances in stem cell biology. In particular, the use of prospective isolation strategies to define stem cell populations has made detailed analysis possible. Stem cells clonally expand to give rise to further differentiated progeny, initiating gene expression programs along the way. The most important function of a stem cell self-renewal, is a unique program. These developmental characteristics of tissue architecture are an excellent model system for the application of gene profiling technology. The major obstacle of this process is obtaining a highly enriched stem cell population to extract enough RNA to generate microarray data. However, the advantage of expression profiling stem cells is that *a prioi* assumptions are not needed to generate data[1].

Recently, the surface protein phenotype of the murine mammary stem cell population was described[2,3]. One group performed microarray analysis of the MRU, or stem cell enriched, fraction of cells and compared the expression profile to further differentiated cells including bipotent progenitors. Their analysis revealed only 4 stem cell specific genes, two of which were involved in fatty acid metabolism. This result was surprising given the amount of unique functions, such as self-renewal, that stem cells perform differently than the populations they were compared against.

We recently further enriched the murine stem cell population by the addition of Thy-1, or CD90, as a marker. In our hands, Thy-1 expression enriched at least 4-fold over the previously defined phenotype (approximately 1 in 250 cells versus 1 in 1100 cells). Therefore, we now had a population of cells with greater frequency of functional stem cells, and we predicted microarray analysis would yield more stem cell specific genes. To that end, we isolated highly enriched stem cell (MaSC), multipotent progenitor (MMPP), early progenitor (MaCFC), myoepithelial (MYO) and luminal epithelial cells (EPI) populations and performed Affymetrix GeneChip 430 2.0 arrays. Our resulting analysis is the subject of this report.

Microarray technology enables researchers to assess global gene expression and has been extensively used in gene discovery, biomarker identification and studies of gene regulation[4-8]. The generation of large data sets based on gene expression has made necessary the need for bioinformatic tools that help identify genes or pathways of interest. There are a number of publicly available tools and resources to aid researchers in functionally annotating gene lists to aid investigation of which genes in their *a priori* gene set correlate with specific biological functions, molecular functions or cellular components[9]. These methods provide an exploratory platform to ask questions about a particular cell type, whether in the context of physiological response of an experimental manipulation or the endogenous workings of a population of cells. The population basis of all gene expression data creates a situation whereby interesting finding must be pursued by more traditional biological experimentation to confirm interactions. Regardless of the need for physical validation of results, bioinformatic tools have become a powerful tool to both create new hypotheses and help understand results from complex

biological experiments. Used as a discovery tool, for example the analysis of microarray results, bioinformatics can direct research in an empirical data-driven methodology.

## MATERIALS AND METHODS

### Animals

All animals used in the study were C57Bl/6 mice that were maintained at the Stanford Animal Facility in accordance with the guidelines of both Institutional Animal Care Use Committees.

### Microarrays

Double sorted cells were isolated as described in chapter 1. Freshly sorted cells were immediately centrifuged at 5000 rpm in low attachment Eppendorf tubes. For replicate, the fat pads from 10 mice were pooled together and processed as a single sample. We peformed three replicates for each population. Supernatant was carefully removed and the cell pellet was snap frozen in liquid nitrogen. Pellets were stored in -80C freezer until further processing. RNA was isolated using the mirVana RNA isolation kit (Ambion) according to manufacturer's instructions. RNA was then processed for microarray hybridization by the Stanford Protein and Nucleic Acids core facility. RNA was applied to Affymetrix GeneChip Mouse Genome 430 2.0 oligonucleotide arrays. The resulting CEL images were then processed using the TIGR TM4[10,11] suite of bioinformatic software as described below.

**Gene expression analysis**

Arrays were pre-processed using robust multichip average (RMA) normalization across all samples. For hierarchical clustering, MIDAS (TIGR, TM4 suite) processed arrays were input into Cluster 3.0. Only probes that expressed at least 6.229 (in $Log_2$ space) in any 2 of the arrays and differed by greater than 2-fold (absolute value) in any two arrays were included in the analysis. Data was then clustered based on Pearson uncentered distance similarity using average linkage. Output CDT data was visualized in TreeView 1.1.3. For comparison of two populations, all arrays corresponding to each population were further processed together, but separately from the entire original set. For population comparison testing, arrays were processed in Cluster 3.0 to remove probes that expressed 100 expression units of background noise, and probes whose value did not differ by at least 4-fold in any two arrays. Probes intensities were then median centered across all arrays as a set. A permutation based T-TEST was then performed using the Multiple Experiment Viewer (TIGR, TM4 suite) software with the following parameters: Welch approximation for unequal variances, all unique permutations of array samples, p-value cutoff of 0.01, standard Bonferroni correction, and Pearson Correlation metric for distance in clustering both significant genes and arrays. The significant genes were then processed in Excel 2007 (Microsoft) and average expression of each probe in each population was calculated. Probes whose average expression in each population was at least 2-fold different were retained. This list was then clustered again in MeV, and gene order was saved to a separate file. Genes that had a higher expression in a given population was color red and lower expression was colored green. Scale intensities were

changed to -2 (green) to +2 (red) in the heat map image. Red genes were then used to create the gene list of upregulated genes as noted in the Results section. Principal Components Analysis was performed according to MeV manual instructions.

**GSEA**

GSEA was performed as previously described. All analyses were performed through the GSEA website of the BROAD Institute (www.broad.mit.edu/gsea/), version 2.04. MIDAS RMA normalized expression data set including all mammary epithelial populations were used. Parameters that deviated from default setting were gene set permutation, and exclude minimum gene sets contained less than 1 gene. Median of probe method was used to collapse multiple probe sets for an individual gene. Chip platform file was obtained from Affymetrix for the 430 2.0 array used in all analyses. Criteria for significant pathways was FDR<0.05 since only a small number of samples was tested for each population.

**Real-time PCR for expression of individual genes**

Sorted cell populations were collected in staining media directly and then centrifuged at 5000 rpm for 5 min at 4°C. Supernatant was then carefully removed from the cell pellet which was immediately frozen in liquid nitrogen and stored at -80°C until RNA extraction. RNA was extracted from frozen cell pellets by Trizol. RNA was resuspended in MilliQ $H_20$. RNA was then converted to cDNA using the Superscript III Reverse Transcriptase system (Invitrogen). QPCR was then performed on fresh cDNA using SYBR Green PCR Master Mix (Applied Biosystems) according to the

manufacturer's instructions. All primers used in these studies have been previously described[12,13]. GAPDH expression was used as the reference for all experiments.

**Immunocytochemistry**

Whole mammary glands were fixed in formalin and embedded in paraffin. Three-micrometer sections were dewaxed, hydrated and microwaved for 10 min in Tris-EDTA (0.01-0.001 M; pH 9) for antigen retrieval. Tissue sections were incubated o/n at 4°C with primary antibodies in TBS+1%BSA. Antibodies were CK6 (Covance), CK5 (Covance), Troma-I (DSHB) and Troma-III (DSHB). Sections were then incubated with anti-rat –Alexa A488 antibody (Invitrogen) with anti-rabbit A594 antibody (Invitrogen) for 30 min at room temperature. Secondary antibodies were applied at 1:200 dilution. Samples were stained with DAPI and mounted in ProLong before pictures were taken. All images were produced with a Leica microscope and Image Pro Software. Images were opened in Photoshop CS2 (Adobe) to convert the images from 16-bit RGB to 8-bit RGB, but no other processing was applied.

***Ex vivo* protein analysis**

Sorted cell populations were stained using components of the BrdU Flow Staining Kit (BD). Cells were resuspended in 150ul of BD Cytofix/cytoperm buffer and fixed for 20 min at room temperature, then washed with BD Perm/wash buffer and resuspended in BD Cytoperm plus buffer and incubated on ice for 10 min. Then the cells were washed again and resuspended in BD Cytofix/cytoperm buffer and incubated on ice for 5 min.

Cells were then washed and resuspended in BD Perm/wash buffer. Primary antibodies used were CK8 (Troma-1, DSHB), CK19 (Troma-III, DSHB), ESA (G8.8, DSHB), or CK14 (Covance). After a 20 minute staining incubation on ice, cells were washed with BD Perm/wash buffer, and then incubated with either donkey anti-rat IgG Alexa 488 (Invitrogen) or goat anti-rabbit IgG Alexa 488 (Invitrogen) for 20 minutes. Cells were then washed once more and resuspended in 200 ul of BD Perm/wash buffer before analysis by flow cytometry.

### *In vitro* colony forming assay

Co-culture colony forming assays were performed as previously described[3]. Briefly, irradiated NIH3T3 cells are plated into 24 well tissue culture plates (Costar) in Epi-Cult media plus 5% FBS (Stem Cell Technologies). Sorted cells were then plated and media was changed to serum free media 24 hrs later. After 7 days, colonies were stained for immunofluorescence. Colonies were stained with the BD BrdU FITC Flow Staining Kit (BD) with Troma-1, Troma-III, ESA and CK14. Cells were also stained with DAPI to mark nuclei prior to imaging. All images were produced with a Leica microscope and Image Pro Software. Images were opened in Photoshop CS2 (Adobe) to convert the images from 16-bit RGB to 8-bit RGB, but no other processing was applied.

## RESULTS

Bioinformatic analysis of gene expression data may be used to answer a particular hypothesis or to generate new ones. There are vastly different strategies to accomplish these goals; therefore a clearly defined endpoint for analyzing microarray data is

necessary. We were specifically interested to use microarray technology to investigate stem cell specific gene sets. Our previous work showed that MaSC cells from nulliparous adult murine mammary glands are enriched for stem cells. In addition, the multipotent progenitor cells (defined as MMPP) had diminished self-renewal. This work added to the existing cellular hierarchy in the mammary gland. Using these newly defined populations, we sought to define genes that were specifically expressed by the MaSC population. Gene expression profiling of isolated mammary populations provided an excellent method to accomplish this goal, as we could also use the data to define genes specific for further differentiated cells.

We began by double sorting MaSC, MMPP, MaCFC, MYO and EPI populations to ensure the best enrichment for each population. Due to the limited number of stem cells that are present in fat pads of an adult mouse, the mammary tissue from multiple wild type C57BL/6 mice, all 6-8 weeks old, were pooled and processed together for a single experiment to minimize variability. Due to the time limitations of processing and FACS, multiple sorts were pooled to achieve the numbers of cells necessary for stem cell and multipotent progenitor arrays. We justify this strategy since these genetically identical mice were housed and raised together, and thus were subjected to similar environmental effects (i.e. food, water, etc). The pooled cell numbers from our sorts that were used to generate RNA for each array is listed in Table 2.1. For each replicate array (A, B and C) of the MaSC and MMPP populations, the fat pads from 10 mice were pooled and processed as a single sample for sorting. The cells from multiple sorts=ing days were pooled together to obtain approximately 10,000 cells per array. All mice used were the inbred C57BL/6 strain that were between 6-8 weeks of age. The amount of cells

we used for the MaSC arrays were approximately two to three times greater than previously published microarrays[3], providing increased signal intensity that we predicted would help to isolate differentially expressed genes. RNA was extracted from the cell populations and hybridized to Affymetrix Mouse 430.2 arrays which have probes that correlate to most genes in the mouse genome. The resulting CEL files were then processed using various programs from the TIGR bioinformatic suite of software[14]. Arrays were normalized and background noise removed using MIDAS.

**Mammary stem cells have a basal cell phenotype according to their cytokeratin signature**

Our array analysis began with unsupervised hierarchical clustering of all arrays based on Pearson correlation using the average distance metric (Figure 2.1). The clustering analysis indicated which arrays are most similar in their overall gene expression, with the resulting familial relationships displayed as dendrogram information[15]. This clustering method indicated MaSCs are most similar to the MYOs, corroborating ours and others[16] previous histological data that Thy-1 localizes near the basal lamina of mature ducts, in the same region as myoepithelial cells. One array, MaSC B, had a closer relationship to the MYO B and MMPP A arrays than to the MaSC A and MaSC C arrays. Interrogation of probe level data revealed MaSC B had the weakest expression of Thy-1, although it was upregulated compared to any of the MMPP arrays. We performed Principle Components Analysis (PCA) to assess the similarity of the arrays (data not shown). This analysis also showed the MaSC B array segregated closer to MMPP arrays

than the other MaSC arrays. We chose to include the MaSC B array in future analyses and comparisons because probe level information for certain known stem cell marker genes (i.e. Thy-1, CD24, etc.) was most similar to the other stem cell arrays. Therefore, we could not determine if MaSC B was an outlier array or a better indicator of real stem cell gene expression without performing more replicates. Further inspection of the dendrogram also showed the MaCFC arrays were most similar to the EPI arrays. Taken together, the clustering results suggested the stem and multipotent progenitor cells have a phenotype similar to myoepithelial cells and the further differentiated MaCFC progenitors have a luminal signature. These conclusions supported our previous immunohistochemistry findings that Thy-1 localizes near the basal lamina and Thy-1 positive stem cells, like myoepithelial cells, are poor colony forming cells *in vitro*.

We next examined differentially expressed cytokeratins (KRT) to make sure our arrays were consistent with the literature, has these proteins have long been used to distinguish the variety of cell populations in the mammary system[17]. Myoepithelial cells express Krt5 and Krt14; both proteins are found on the same intermediate filaments within a particular cell. Luminal epithelial cells express Krt7, Krt8, Krt18 and Krt19. As opposed to other luminal cytokeratins, KRT19 is thought to be expressed on only a subset of luminal cells. In human mammary ducts, KRT19 is expressed on most luminal epithelial cells except during pregnancy, where its expression turns heterogeneous and is a marker of secretory alveolar cells[18]. We performed an analysis to look at cytokeratin expression in the different mammary populations to confirm known results from the literature (Figure 2.2). We took the average expression from each probe from all replicates of a particular population's arrays as representative data.

Our results showed Krt8, Krt18, Krt19 and Krt7 were upregulated in the MaCFC and EPI populations, consistent with previously published results. In addition we detected Krt6 specifically in the MaCFC progenitor population. Krt6 was previously found dispensable for mammary development, although it is expressed in the lobules of the developing mammary gland[19]. This report also showed Krt6 expression is rare in the mature mammary gland. Our data suggests Krt6 may be a specific marker of MaCFC progenitor cells in mammary epithelium, and we expect these cells to be found in the luminal cell layer based on our hierarchical clustering analysis. This hypothesis is also supported by a previous study in which mice that had mutant C/EBPβ in mammary glands caused Krt6 to be expressed only in luminal cells[20]. Since our results were consistent with previously published reports of cytokeratin expression, lending support to the validity of the information gained from our array analysis. In addition, there may be another Krt6 negative progenitor population that has the same differentiation potential as MaCFC cells.

To confirm the myoepithelial phenotype of the MaSC and MMPP cells we analyzed the overall cytokeratin signature in those populations (Figure 3.2). Again, consistent with previous reports[21-24], those populations expressed Krt14 and Krt5. Of the cytokeratins expressed in the basal populations, there were two groups of proteins: the Type I keratins Krt16, Krt23, Krt14, Krt10, Krt15 and the Type II keratins are Krt17 and Krt5, Krt80, Krt1 and Krt77[25-27]. Since previous data has shown myoepithelial cells express both Krt5 and Krt14, these newly identified keratin molecules may prove useful markers of basal/myoepithelial cells. In addition, Krt17 has been proposed as a basal marker specific for myoepithelial cells[28-31]. Of particular interest was Krt80, which

was expressed by the stem cells and differentiated luminal epithelial cells. There is no previous data on Krt80 in the mammary gland. Krt78, a Type II keratin, was upregulated in the MaSCs compared to the rest of the populations. Krt23 was also an interesting molecule which was expressed by the multipotent progenitors and differentiated myoepithelial cells, but not the stem cells. This cytokeratin may be useful in identifying the stem cells in histological analysis by screening sections for basal cells that are negative. Since the MaSC population expressed Krt80 but not Krt23, our cytokeratin data suggest stem cells retain markers of both luminal and myoepithelial lineages, as has been previously suggested[24,32].

**MaSCs activate a luminal epithelial transcriptional program as they differentiate into MaCFCs**

Since microarray data relies on hybridization kinetics to indicate gene expression, further validation of probe level data is required to confirm results. To that end, we performed semi-quantitative relative real-time PCR using SYBR green using Gapdh as our reference gene as previously described{Asselin-Labat, 2006 #156}. Differentiated luminal and myoepithelial populations are the best described cell types in the mammary gland, so we first began by measuring gene expression of common cytokeratins and the actin component SMAα in isolated Lineage-, MaCFC, EPI and MYO populations (Figure 3.3A). Our data (Figure 3.3, red series) showed that Lineage- cells expressed Krt18 and to a lesser degree Krt19, as well as low expression of SMAα, reflecting the heterogeneity of the mammary epithelial cells. We also looked at cytokeratin expression in both the MYO (Figure 3.3, blue series) and EPI (Figure 3.3, purple series) populations. Consistent

with previous reports, we found MYO cells had high expression of SMAα but we did not observe a similar level Krt14 expression, which is also a myoepithelial marker. However, no luminal epithelial specific genes such as Krt18 were expressed by MYO cells, suggesting the qPCR results accurately reflected the identity of the sorted population. EPI cells showed extremely high expression of Krt18 and Krt19 to a lesser extent. The EPI cells did not express any markers of myoepithelial lineage. When we analyzed MaCFC (Figure 3.3, green series) progenitor cells, we found a similar profile to EPI cells, suggesting these cells had begun differentiating into luminal epithelial cells. This data is consistent with our *in vitro* colony forming data which demonstrated the lack of cell colony forming ability correlated to a more basal/myoepithelial phenotype, whereas the MaCFC and EPI populations had an increased ability to form colonies *in vitro*.

Next, we analyzed the cytokeratin expression of the MaSC, MMPP and MaCFC subpopulations. We observed the MaSC population had a myoepithelial mRNA expression profile, with significant upregulation of SMAα and Krt14. The MMPP population had a similar profile to the MaSC cells, also expressing myoepithelial markers. In both MaSC and MMPP cells, we observed a small amount of Krt18 expression, suggesting these cells may co-express luminal and myoepithelial markers as a reflection of their undifferentiated status since expression of both luminal epithelial and myoepithelial markers has been suggested to mark progenitor cells. We also observed a decrease in the level of Krt14 in MMPPs as opposed to MaSCs, without a noticeable reduction in the profile of SMAα. Our initial observations showed the MaCFC cells had differential Thy-1 expression. We found both Thy-1$^{+}$MaCFC and Thy-1$^{-}$MaCFC cells had a similar expression profile, with high expression of Krt18 a lower amount of Krt19.

We also observed the Thy-1$^{+}$MaCFC population expressed SMAα and Krt14, but the Thy-1$^{-}$MaCFCs did not. This result suggested the Thy-1$^{+}$MaCFCs may be upstream progenitors that have not fully committed to a particular lineage while the Thy-1$^{-}$ MaCFCs may be downstream luminal restricted progenitor cells. We found a very low amount of Krt14 and SMAα was expressed by these progenitor cells, suggesting these cells were differentiating into luminal epithelial cells. The stark contrast of MaSC versus MaCFC cells confirmed our hypothesis of a change in the transcriptional programs active in these two populations. Interestingly, MaCFCs can make both differentiated myoepithelial and luminal epithelial cells, suggesting self-renewal may be confined to basal cells that express myoepithelial makers. We also tested the expression of Krt6 in all cell populations. We failed to find expression of Krt6 in any isolated population. This data contradicted previous literature which suggested Krt6 is a marker of stem and progenitor cells, but was not found in terminally differatiated cells[23]. This was the only gene whose expression from our microarray data could not be reproduced by real-time PCR. This negative result may have been due to poor primer design since we did not detect expression in any of the samples tested or our assay may not have been sensitive enough to detect expression of this gene.

To further validate our microarray analysis, we chose to investigate the expression of *p21$^{cip1/waf1}$* and *Tbx3* in progenitor populations. These two genes have been suggested to have roles in mammary morphogenesis. *p21$^{cip1/waf1}$* has been hypothesized to maintain stem cell quiescence in the skin system[33,34], and thus may be considered a putative stem cell marker. In integrin β1, or CD29, mutant murine mammary glands luminal cells had reduced proliferation[35] which correlated to an increase in the expression of

*p21*$^{cip1/waf1}$. The authors suggested p21$^{CIP1/WAF1}$ participated in the repression of luminal cell proliferation in the mammary gland. A previous study showed exogenous expression of *TBX3* resulted in the reduced expression of the CDK inhibitor *p21* $^{CIP1/WAF1}$ [36]. *TBX3*, a gene often mutated in Ulnar-Mammary Syndrome (UMS), is a transcription factor that is part of the T-domain containing family. Haploinsufficient UMS patients experience a congenital loss or reduction of mammary epithelium[37]. Homozygous *Tbx3* null mice are in part characterized by a loss of mammary rudiments by E13[38]. TBX3 has been shown to repress the transcription of the ARF protein in non-mammary cells[39]. In addition, overexpression of *Tbx3* in mouse embryonic fibroblasts leads to their immortalization[40,41]. These data suggested a role for *p21*$^{cip/waf1}$ and *Tbx3* in stem cells or early progenitors.

We began by looking at all probes for both genes (Figure 3.4, top panel). Althoguh we were not able to statistically distingusigh these genes in our global array analysis, probe level information showed *p21*$^{cip1/waf1}$, commonly known as *Cdkn1a*, had high expression in the MaSCs and declined as the cells differentiated into MMPP and MaCFC. This expression pattern was opposite of *Tbx3*. The MaSCs specifically did not express *Tbx3*, but its expression was detected by one probe in MMPP, and many of the probes in MaCFCs. This data was interesting as Tbx3 appears to discriminate between the stem and multipotent progenitor cells. We validated the expression of *p21*$^{cip1/waf1}$ in the stem and progenitor populations by real-time PCR. We found *p21*$^{cip1/waf1}$ was specifically expressed in the MaSC population and absent in the MMPP population as well as both Thy-1 MaCFC subpopulations (Figure 3.4, blue series). There are two splice variants known for p21cip1/waf1, and we designed our primers to target the exons that

are common between both, thus we cannot discriminate if there is a more prevalent isoform. *Tbx3* was not expressed in MaSCs but was expressed by MMPP and Thy-1$^{+}$MaCFCs. When compared to the expression of these genes in Lineage$^{-}$ cells, we observed enrichment in the described populations, showing the specificity of our results. We hypothesize *p21$^{cip1/waf1}$* may function to maintain stem cell quiescence in murine mammary stem cells. These results also suggest *Tbx3* affects the progenitor population(s) of the murine mammary gland in a *p21$^{cip1/waf1}$* dependent manner, but not the stem cell population.

Taken together, our results show stem, multipotent progenitor and myoepithelial cells had increased expression of SMAα and Krt14 relative to the other populations. MaCFC progenitor and differentiated luminal epithelial cells had high expression of Krt18 and to a lesser extent Krt19. Consistently with our cytokeratin cluster microarray analysis, MaCFC progenitor cells expressed both cytokeratins 18 and 19, confirming their luminal identity. MaSCs and MMPPs had a basal/myoepithelial phenotype. In addition, array expression analysis of *Tbx3* and *p21$^{cip1/waf1}$* was confirmed by real-time PCR. Extrapolation of these data may implicate multipotent progenitor defects rather than stem cell anomalies as the source of congenital UMS in human patients.

**Mammary stem cells express luminal epithelial proteins upon differentiation into progenitors**

After we validated the microarray expression data, we investigated how certain keratin proteins are expressed in isolated mammary populations. Post-transcriptional protein regulation can dramatically alter conclusions based on expression data alone. *Ex*

*vivo* analysis of isolated cells in combination with immunofluorescence staining had been previously shown to describe the keratin profile of MRU and MaCFC cells[3]. We used an adaptation of this approach and subjected the cells to gentle fixation and permeabilization as previously described[42] to measure the percentage of cells that expressed KRT8, KRT19, ESA and KRT14 in sorted mammary populations (Figure 3.5). Our data is the first report of murine ESA, or epithelial specific antigen. ESA, also known as EpCAM, has proven valuable to isolate epithelial stem and cancer stem cell populations[43-51]. However, the lack of a murine ESA antibody has prevented parallel enrichment strategies in mouse models. We found a rat anti-mouse monoclonal ESA antibody, clone G8.8, from the Developmental Studies Hybridoma Bank (DSHB). This antibody was initially described as a marker of thymic epithelium[52].

As controls, we analyzed the profile of unsorted and Lineage$^-$ (Lin$^-$) cells. Unsorted cells contained a small amount of cells that expressed the luminal epithelial makers KRT8, KRT19 and ESA, but a negligible amount of KRT14 cells. This profile was conserved in Lineage$^-$ cells, where 20% of cells expressed KRT8 and about 28% that expressed KRT19 and ESA. KRT14 was expressed in a smaller percentage of cells, with only 6% of all cells staining positively. These results were expected since removal of blood cells essentially enriches for epithelial cells. We found KRT14 was expressed in about 40% of the MYO population, and the other proteins on only a small fraction of the cells. We found 81% of EPI cells expressed ESA, 63% expressed KRT19 and 38% marked KRT8.

There was no difference in the percentage of cells that expressed lineage identification proteins between MaSCs and MMPPs. Previous reports indicated KRT14

was highly expressed in the $CD24^{med}CD49f^{hi}$ population in mRNA[3]. KRT14 was highly expressed in the MaSC and MMPP populations, with 41% and 52% of cells staining positive, respectively. ESA was found in about 20%, KRT19 in 15%, and KRT8 approximately 5% of both fractions of the MaSC population. Taken together with previous reports and based on the MYO and EPI protein data, we conclude that Thy-1 does not isolate a luminal epithelial population from the $CD24^{med}CD49f^{hi}$ population. When we analyzed the Thy-$1^{+}$MaCFC and Thy-$1^{-}$MaCFC populations, their profiles were also remarkably similar. We found most of the MaCFC population expressed KRT8, KRT19 and ESA proteins. These increases were concomitant with a precipitous decrease in KRT14 expression compared to MRUs, suggesting the differentiation process from MRU to MaCFC induces a luminal epithelial transcriptional program. When we analyzed the difference in the percentage of KRT14 positive cells between the Thy-$1^{+}$MaCFC and Thy-$1^{-}$MaCFC populations we found a significantly lower amount of KRT14 positive cells in the Thy-$1^{+}$MaCFC population versus the Thy-$1^{-}$MaCFC population (Student's T-test, 2-tailed, P=0.029). Since the Thy-$1^{-}$MaCFC cells are more similar in expression profile to the EPI cells than the Thy-$1^{+}$MaCFC cells, we suggest these cells may be subset of the MaCFC population destined to become differentiated luminal epithelial cells. As a caveat, we did not observe a decrease of KRT19 expression, which may indicate these cells are not fully differentiated.

Our protein expression data confirmed our results from real-time PCR and microarray gene expression. Although the MaSC and MMPP populations expressed markers of both luminal and myoepithelial lineages, we found these populations had a high amount of KRT14 protein expression. In contrast to our single gene expression data,

we observed KRT19 expression in the MaSC and MMPP populations. We also report the ESA is expressed by all of the mammary populations analyzed, and most of the EPI and MaCFC cells are $ESA^+$. Using ESA in combination with cancer stem cell markers may help to define the cell of origin in tumors, or assess the transition of basal cells into adenocarcinoma.

**Cultured mammary epithelial cells express luminal and myoepithelial proteins**

To better define how KRT8, KRT14, KRT19 and ESA are expressed, we stained colonies that contained differentiated cells. MaCFC cells were sorted and plated into 3T3 co-culture assay, followed by indirect immunofluorescence staining to detect specific proteins. We show the first *in vitro* data for ESA using the antibody G8.8. Previously, the G8.8 antibody was used to mark subcapsular and medullary thymic epithelial cells[52,53] and also labeled gut, epidermal and some kidney epithelium. We stained cultured mammary epithelial cells and found ESA expressed diffusely in the cytoplasm of luminal epithelial cells, with strong staining at the cell surface and boundaries between these cells (Figure 3.6A). This staining is similar to ESA staining of cultured human mammary cells. KRT19 was detected using the Troma-III antibody, also obtained from the DSHB. Troma-III was also used a marker of differentiating embryonal carcinoma cells[54]. This antibody was found to mark EndoC, a trophectodermal protein, which was identical to murine KRT19[55]. Although KRT19 is thought to label a subset of luminal epithelium, we found only rare cells were labeled with this antibody. These cells were always found in luminal epithelial colonies, but the cells had a small amount of cytoplasm, were rounded, and typically only a few cells were found at the edge of colonies (Figure 3.6B).

Therefore, we conclude that KRT19 expression is down-regulated when MaCFCs differentiate into luminal epithelial cells *in vitro*. KRT14 was detected using a polyclonal rabbit antibody previously described to be useful for mammary lineage identification. Cultured myoepithelial colonies were positive for KRT14 (Figure 3.6C). Troma-1, an antibody shown to mark the endodermal cytoskeletal protein Endo A, was used to mark trophoblast cells, Merkel cells, auditory receptor cells, and differentiating embryonic stem and carcinoma cells[56-59]. The Troma-1 antibody was also found to react with KRT8, a keratin of simple epithelia[60-62]. KRT8 has been used as a marker of normal alveolar and ductal epithelia in mouse mammary tissue[63,64]. When we stained cultured mammary epithelial cells with KRT8, we found luminal epithelial colonies stained positive for KRT8 (Figure 3.6D).

**Immunofluorescence of mature mammary ducts**

*Ex vivo* protein analysis of sorted cell populations and *in vitro* colony staining confirmed our previous keratin expression results. We hypothesized that using keratin molecules on fixed sections of adult mammary tissue would help identify the location of progenitor cells. Immunofluorescence co-staining of KRT8 and KRT14 clearly distinguished the KRT14 positive myoepithelial cells from the KRT8 positive luminal cells (Figure 3.7, top left). We did not observe any yellow staining which would indicate both cytokeratins are simultaneously expressed, as a marker of MaSCs. These cells are rare, and thus screening many more ducts would be required to accurately find them. KRT5 also localized only to the myoepithelial cells (Figure 3.7, middle left). We found Krt6 expression in a small number of KRT8 positive cells in the luminal layer of ducts

(Figure 3.7, bottom left), consistent with our microarray data that Krt6 was expressed in MaCFCs that retained a luminal identity. This result shows MaCFC cells are not a luminal expressing sub-population of basal cells. After an exhaustive literature search, we could not find any histological data of the expression of KRT19 in murine mammary glands. We present the first staining of Krt19 in the murine ductal epithelium (Figure 3.7, top right). Consistent with previous data in human breast, Krt19 marked a subset of luminal murine cells. An attractive speculation would be the Krt19 marks cells that have alveolar secretory differentiation potential, although further staining in pregnant mice or lineage tracing experiments are needed to test this hypothesis. Our results confirm our microarray results which accurately predict the identity of a given mammary epithelial population based on cytokeratin expression. We were able to identify KRT6 and KRT19 luminal subpopulations. Furthermore, our data suggests newly identified cytokeratin molecules such as Krt80 and Krt23 may also be useful histological markers to distinguish stem and progenitor cells from further differentiated cells.

**MaSC cells have activated molecular functions not found in MYO cells**

Based on our clustering results that the MaSCs share the closest relationship to the MYO cells, we went on to further investigate how our microarray data could identify differences at the cellular level. This approach was advantageous as MYO cells were essentially indistinguishable in histological analysis from MaSC cells. We used Gene Set Enrichment Analysis (GSEA) in which a previously defined set of genes or various sets of genes may be used to identify specific genes or pathways that are preferentially

enriched in a certain population of microarrays[65]. We show the molecular functions that were enriched in the stem cell population using a FDR q-value<0.05 as our cutoff (usually this value is kept at 0.25, but the small number of arrays used in the analysis necessitated a more stringent metric to eliminate spurious results). Our data indicates MaSCs have genes involved in neurotransmitter signaling to be highly enriched (Figure 3.8). Not surprisingly, ion channel pathway members were also enriched, which are known to mediate many neurological functions[66]. In addition, metabolic pathways such as acetylcholine binding, peptide receptor activity and inositol phosphodiesterase activity were also enriched. This analysis suggested stem cells may have different metabolic processes than myoepithelial cells. Also, this analysis provides a hypothesis generating platform in that genes that are "core enriched" may be exploited to specifically module stem cell activity.

**Genes upregulated in MaSCs versus MYO cells**

Since our GSEA results suggested intracellular differences between MaSCs and MYO cells, we performed a bioinformatic screen for genes that were upregulated in MaSCs. Our results were based on a permuted T-Test followed by removal of probes whose average expression was not greater than 2-fold different in the stem and myoepithelial arrays. This analysis resulted in a curated list of 71 unique genes, the first report of a stringent differential gene analysis that reveals such a large set of stem cell specific genes (Figure 3.9). Functional annotation was necessary to investigate which pathways were identified by upregulation of these genes. We chose to process this list

through the DAVID online gene ontology database[67], asking which biological processes these genes enriched for. Intriguingly, five of these genes were involved in cell motility. One of these genes, Zeb2, has been implicated in epithelial to mesenchymal transition[68-70]. Another of these genes, Nr2f2, is an orphan nuclear receptor that has recently been shown to affect embryo attachment and uterine decidualization through a BMP2 specific signaling mechanism[71]. The authors also suggest this effect was progesterone mediated.

The MaSC specific genes also enriched for cellular calcium and other ion homeostasis, correlating with our GSEA results that calcium and potassium voltage regulate channels were also differentially expressed by the stem cells. Another interesting gene was Wnt8b, whose ortholog been shown to be expressed by some human breast cancer cell lines[72], but has been better studied in the hypothalamus where it is required for neurogenesis through a Lef1 dependent mechanism{Lee, 2006 #7550}. Intriguingly, in hepatic progenitors derived from human ES cells the expression of WNT8B was due to POU51(OCT3/OCT4) and GATA3[73]. These results are important in light of a recent study in with a Gata3 null mutation under Krt14 promoter expression was found to completely abolish mammary rudiment development, therefore preventing any ductal growth post-natally[74]. This may implicate Wnt8b as a downstream effector of Gata3 directed transcriptional activation in a mammary specific context, and also sets up a role for Oct3/oct4 in mammary morphogenesis. Interestingly, WNT8B although upregulated in gastric cancer, pancreatic cancer, colorectal cancer, breast cancer and embryonal tumors it is not found in the normal tissue counterparts of these tumors by Northern blot analysis[72,75]. If indeed Wnt8b is a stem cell specific canonical Wnt, Northern blot

analysis would not have been sensitive enough to detect its presence. Continuing our analysis, there were also four genes that were associated with central nervous system development: Nr2f2, Zeb2, Agtpbp1 and Ngfr. Although it is tempting to speculate the role of these genes, further work is needed to functionally validate how these genes affect mammary morphogenesis.

**Genes upregulated in MaSCs versus MMPPs**

We next performed a similar analysis for genes specifically upregulated in stem cells versus multipotent progenitor cells (Figure 3.10), as we are the first group to describe the population. In a similar screen for genes that were both differentially expressed and upregulated in the stem cells, we obtained 88 unique genes. We obtained more genes in this analysis than in previous ones due to the increased dissimilarity between the stem and multipotent progenitor cells, consistent with our clustering analysis. We did not find any overlap with the previously published 4 stem cell specific genes. Again, we processed our gene list through the DAVID functional ontology database to assess which biological processes were specific to the stem cell population. Interestingly, cell motility, cell migration, cell-cell adhesion and nervous system development were among the most significant results. These results suggested the stem cell population may be motile in the mammary system, perhaps responsible for the invasion process that characterizes mammary development. Indeed, many groups have shown circumstantial evidence of an undifferentiated population cells at the leading edge of lobules where there is direct stromal contact[76-78]. The presumed motility of the

stem cells may provide them with the unique ability to traverse the mammary system to repopulate areas during hormonal stimulation. Again, neuronal development pathways also were enriched for by the stem cell specific genes, which may be a reflection of innervation of the breast. In addition, metal ion metabolism also was significant, a property that defined the MaSCs versus both MYO and MMPP cells. Not surprising given that MaSCs are very similar myoepithelial, the two genes Gucy1a3 and Atp1a2 are involved in the regulation of smooth muscle contraction. These results were corroborated by submitting the gene list to the term enrichment algorithm of the AmiGO online gene ontology program[79]. These results suggested a key difference between the stem and multipotent progenitor cells was locomotion..

**Gene signature of MaSCs**

Given the novel pathways and genes identified by our previous comparisons of genes specifically upregulated in the MaSCs versus MMPP and MYO cells, we sought to expand our analysis to identify genes that were only upregulated in the MaSCs and no other populations. To that end, we identified genes using the permuted T-Test with stringent criteria to identify genes and selected those genes that were upregulated in the MaSC population compared to each individual population. We then pooled those gene list together and selected only those genes that were only expressed in the MaSC arrays. Our analysis revealed 17 unique genes and one predicted gene. The genes we found are as follows: Abca8a, Aspa, Atp1a2, Atp1b2, Cadm2, Ccl11, Edg3, Inpp4b, Kcna1, Nr2f1, Ngfr, Pdzd2, Rarres2, Slc35f1, Tmod2, Ttyh1 and Zeb2. The sole predicted gene was

ENSMUSG00000074335, which shares similarities with the voltage regulated channel family of genes. Complete gene ontological analysis of this gene set is presented in Figure 3.11. We performed the most extensive ontological analysis on this gene list, as it represented the most interesting genes. We looked at all known biological processes and molecular functions from both the gene ontology of the GO[80] and the PANTHER[81] databases. When investigating which biological processes were activated specifically by the MaSC population, pathways for ion homeostasis, adhesion, motility, migration and cytokine signaling are enriched. In addition, cell surface receptor mediated signal transduction is also active. In particular, the most enriched pathways were cytosolic calcium ion homeostasis and potassium ion transport. A more in-depth perspective of the function of stem cells was obtained by analyzing our gene list for molecular functions. The largest enrichment was found in sodium:potassium exchanging ATPase activity. A number of ion signaling pathways coupled to ATPase production are also enriched as well as transmembrane transporter activity. These enrichment results by themselves cannot be used to provide functional evidence of stem cell behavior. However, reviewing which general mechanisms are stem cell specific paints a broad picture of stem cell behavior.

## DISCUSSION

Microarray expression data of isolated cell populations provides a powerful method to generate hypotheses about the function of cells type. In addition, gene

profiling is useful to indentify new pathways that may be exploited to control or direct cells' behavior. In our study, we sought to identify genes that were specifically expressed by murine mammary stem cells. Using our recent phenotypic description Thy-1$^{-}$ CD24$^{med}$CD49f$^{hi}$ to define a population enriched for functional stem cells, we generated expression data that was used to interrogate the lineage identity of these cells as well as the specific molecular functions that makes them different from other mammary cells. We compared the MaSCs to MYO and MMPP cells as these were the populations that shared the most similarity. Therefore, we reasoned only significant characteristics of MaSCs would be observed.

This study yielded a number of interesting results, many of which had not been shown until this report. MaSC cells had a basal phenotype according to the gene expression profiling of keratin molecules. Although MMPP cells shared this phenotype, we discovered the expression of two keratins, Krt23 and Krt80, may help to distinguish the MaSC cells. We validated the microarray expression data with real-time PCR, confirming the MaSC cells induced a luminal transcriptional program as they differentiate into MaCFC progenitor cells. We also showed p21$^{cip1/waf1}$ was expressed only in the MaSC cells, but Tbx3 was expressed in the immediate downstream progenitor MMPP cells. This data also demonstrated that MaSC cells have specific genes that distinguish them from all other known mammary populations. To further confirm our microarray results, we used a novel *ex vivo* protein analysis that demonstrated the basal identity of the stem cells, but also that they express proteins associated with both luminal and myoepithelial lineages. One of these proteins, ESA, had not been described before this report. To better characterize ESA in the mammary system, we stained cultured cells

and showed the staining profile was extremely similar to human ESA[82]. In addition, we found the putative alveolar progenitor marker Krt19 was significantly downregulated when mammary cells are grown in culture. We speculate that lactogenic conditions may stimulate production of KRT19. Since many of the keratin molecules have been used to mark specific mammary compartments, we also performed immunocytochemistry on paraffin sections of adult virgin mammary glands. We showed KRT6 is a subset on luminal cells, perhaps the MaCFC population. Also, we demonstrate KRT19 marks a subset of luminal cells, similar to human breast epithelium.

We reported the first detailed gene expression analysis of highly enriched mammary stem and multipotent progenitor populations. Also, we are the first group to show functional annotation of murine mammary stem cell specific genes, providing a new profile of adult stem cells. In general, we found stem cells preferentially upregulate pathways responsible for ion homeostasis. Furthermore, the stem cells have initiated transcriptional programs for general development and organogenesis, as well as motility and migration. Although the data are a complex network of individual observations, we propose a model in which all of the information may be integrated. Stem cells are incredibly powerful components of the mammary gland, with the ability to regenerate the entire organ or repopulate certain areas. Many groups have hypothesized that stem cells are under hormonal regulation, directed to differentiate or self-renew based on systemic and/or local steroid signaling. Steroid signaling processes did not appear as activated in the stem cells, nor were hormonal receptor signaling pathways involved. Nevertheless, given their importance in the mammary tissue, stem cells must respond to environmental cues. The data suggests stem cells are particularly sensitive to calcium, potassium and

sodium ion signaling. These metal ions are common signaling components, found in many cell types. We hypothesize the stem cells have these signaling components present on their cell surface, such that they are primed for a metal ion signal at any time. This interpretation may be extrapolated to suggest the stem cells respond to ion gradients found in the local microenvironment. An attractive speculation is in response to ion signaling, the stem cells activate energy producing pathways first to gear up for purpose-driven transcriptional activation. An example of this may be the upregulation of ion receptor-mediated phosphorylation. Phosphorylation of target proteins is a well studied mechanism of transcription activation that usually activates a specific set of proteins. This creates a temporal picture of stem cell activation, where these cells are constantly listening for ion based signals. Upon receiving an appropriate signal, the stem cells then first make ATP energy and then activate particular signal transduction mechanisms. Another interesting attribute of the stem cells is the upregulation of genes involved in motility and migration. The mature mammary gland is thought to be relatively static, and thus our results show the first evidence that the stem cells may actively move during morphogenesis. These results may help to explain how ducts invade into the stroma, as it has been observed that there are undifferentiated cells in the caps of the invading lobules and in the epithelial cell layers of the body cells. This provides further evidence for the localization of the stem cells within the ductal epithelium. Alternatively, the stem cells may move in response to local signaling in the ductal tree to repopulate a particular area or expand at a branching point. Adhesion pathways are also enriched by the stem cells genes, and given their basal phenotype this suggests that local niche interactions and ion signaling may be the two most influential systems that initially distinguish stem cell from

more differentiated progeny. Taken together, our analysis provided new insights into both the systems that govern stem cells as well as a temporal model for their intracellular signaling activation.

Since both GSEA and ontological analyses reveal such a high association of the stem cells with neural pathways, there are two major possibilities that may explain this data. The first is the phenotypic Thy-$1^{+}$CD$24^{med}$CD$49f^{hi}$ population, although functionally enriched for mammary stem cells, may also contain neural progenitor or differentiated cells. Alternatively, our stem cells may be secreting factors and directing nerve cell growth in support of mammary morphogenesis. Further experiments using neural specific markers such as glial fibrillary acidic protein may be useful to distinguish between these possibilities. There have been numerous studies that have shown Thy-1 expression on astrocytes[83-87] and a recent study in bone marrow cultures suggested Thy-1 accelerated neurite outgrowths[88].

## TABLES

**Table 3.1: Cell numbers of double sorted mammary populations used for microarrays.**

| | Cell population | | | | |
|---|---|---|---|---|---|
| **Samples** | **MaSC** | **MMPP** | **MaCFC** | **MYO** | **EPI** |
| A series | 9073 | 69657 | 90306 | 240103 | 395660 |
| B Series | 32912 | 46809 | 119422 | 207562 | 213152 |
| C Series | 10000 | 49013 | 198796 | 96462 | 112297 |

**FIGURES**

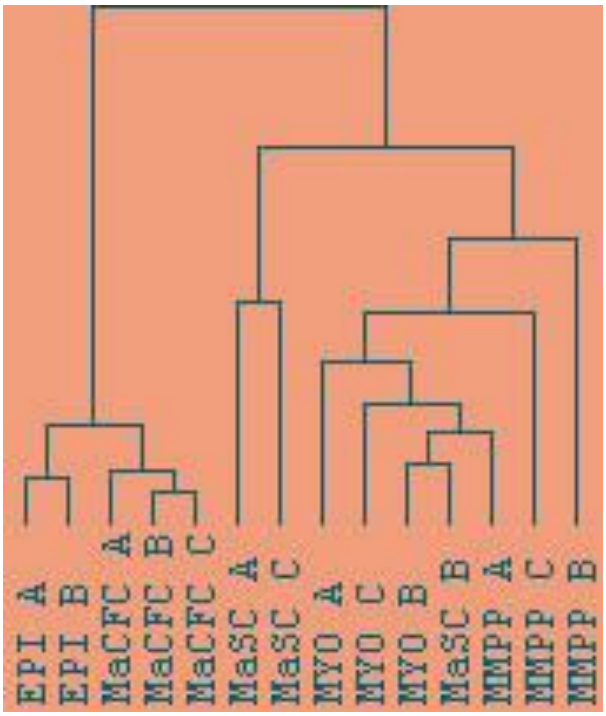



**Figure 3.1: Unsupervised hierarchical clustering of microarrays from mammary epithelial populations.**

RMA normalized microarray data from freshly isolated C57BL/6 mammary cells. The expression data was processed to remove noise and only probe sets that varied by two fold or more were retained. Pearson correlation (uncentered) similarity based on average linkage was used as the clustering distance metric. Luminal, $CD24^{med}CD49f^{-}$ cells, MaCFC, $CD24^{hi}CD49f^{med}$ cells, Stem Cell, $Thy\text{-}1^{+}CD24^{med}CD49f^{hi}$ cells, Myoepithelial, $CD24^{med}CD49f^{low}$ cells, Multipotent Progenitor, $Thy\text{-}1^{-}CD24^{med}CD49f^{hi}$ cells. Each array represents an independent replicate. Data suggests mammary stem cells and multipotent progenitors are most similar to differentiated myoepithelial cells, whereas MaCFC progenitors are most similar to differentiated luminal epithelial cells.

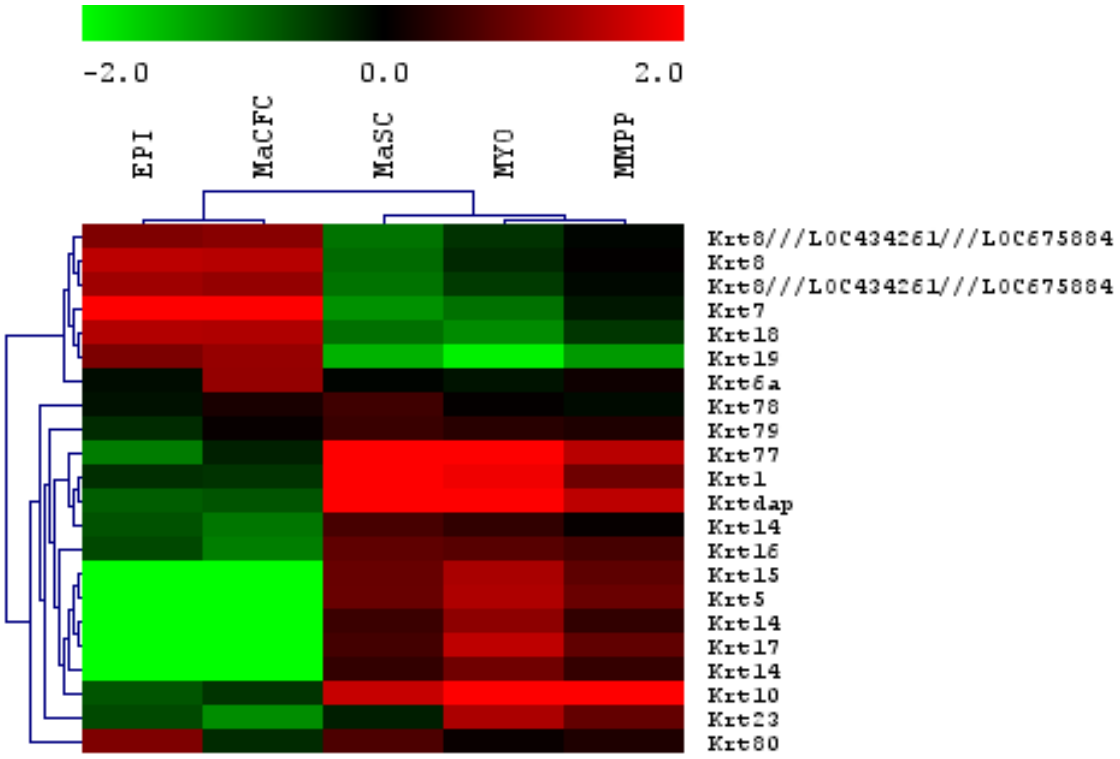



**Figure 3.2: Clustering analysis on average gene expression of cytokeratin molecules from mammary epithelial microarray analysis.**

The most differentially expressed cytokeratins across all populations analyzed are shown. Data demonstrated MaSCs, MMPP and MYO share similar keratin signatures, suggesting a basal phenotype. MaCFC and EPI arrays had luminal epithelial keratin expression signatures.

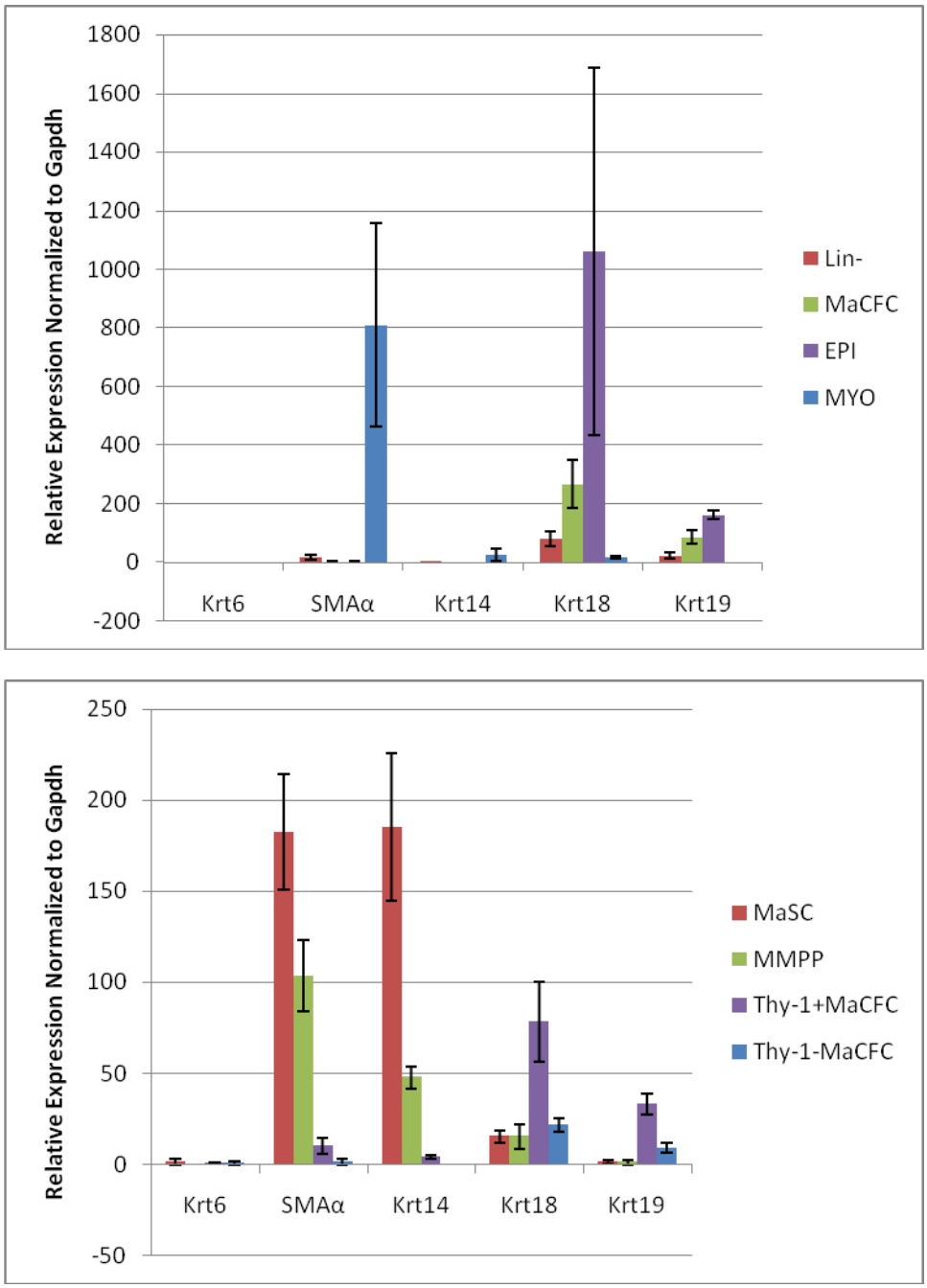



**Figure 3.3: Real-time PCR validation of microarray gene expression analysis.**

Isolated mammary populations were analyzed for the expression of single genes using SYBR green real-time PCR. Top panel, differentiated populations. Lower panel, progenitor populations. All relative expression data was normalized to the expression of *Gapdh* as the control gene. Data represents average of 3 independent experiments, ±S.D of the mean.

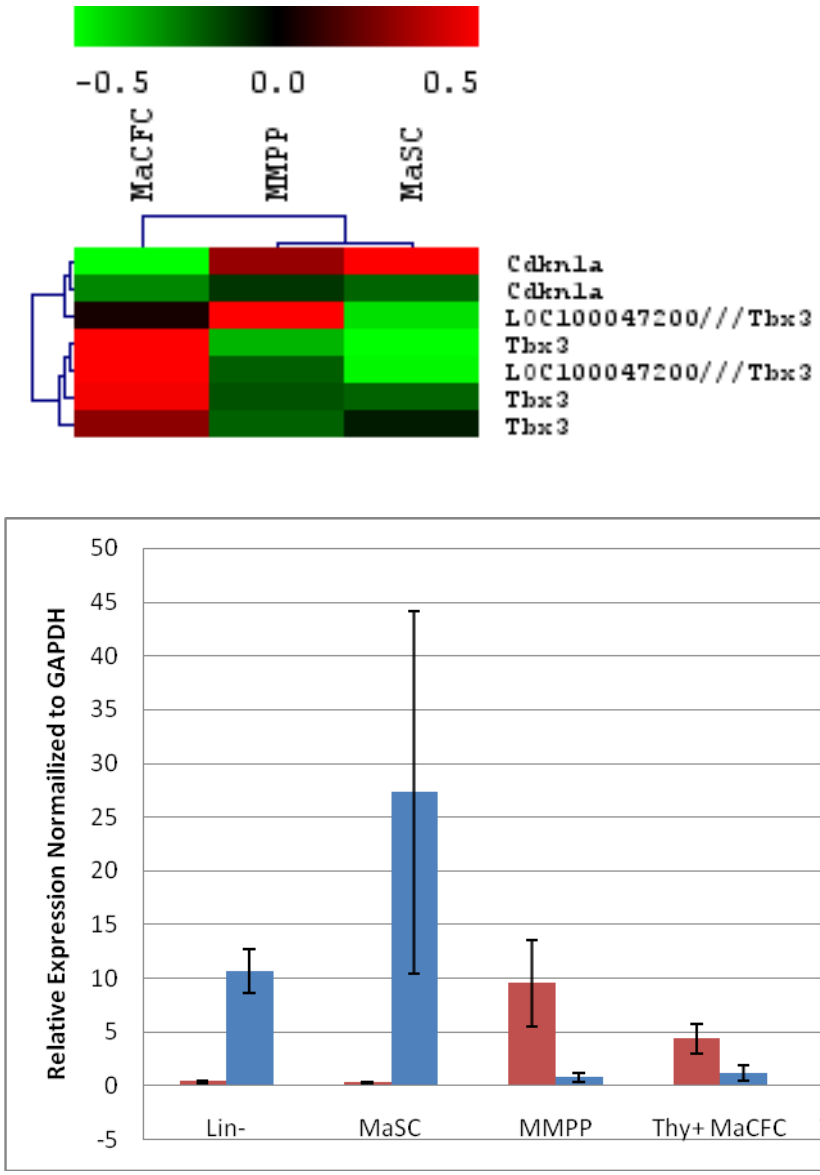



**Figure 3.4: Relative quantitation of *Tbx3* and *p21*$^{cip1/waf1}$ in stem and progenitor populations.**

Top panel, microarray expression cluster of all probes specific for *Cdkn1a* (*p21*$^{cip1/waf1}$) and *Tbx3*. Array and gene clustering based on Pearson centered correlation with average distance metric. Bottom panel, real-time PCR for the same genes in progenitor populations normalized to *Gapdh* as the control gene. Blue series are *p21*$^{cip1/waf1}$ and red series are *Tbx3* data. Data represents the average expression from 3 independent experiments, ±S.D. of the mean.

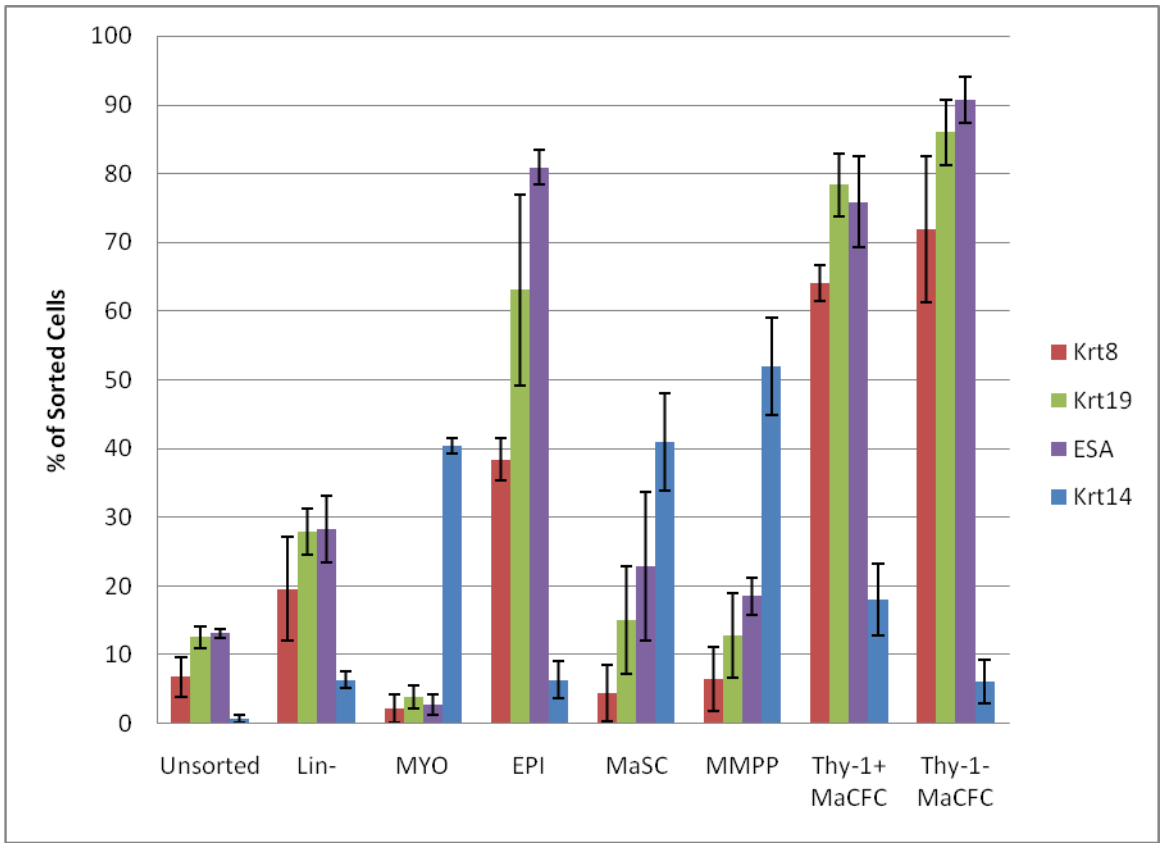



**Figure 3.5: *Ex vivo* protein analysis of sorted mammary cell populations.**

Double sorted cells from each respective population were stained for listed proteins. Data shows MaSC and MMPP cells have similar profiles. Thy-1 does not isolate a subpopulation from MaCFC cells that have a basal phenotype. Results are the average of 4 independent experiments, ±S.D. of the mean.

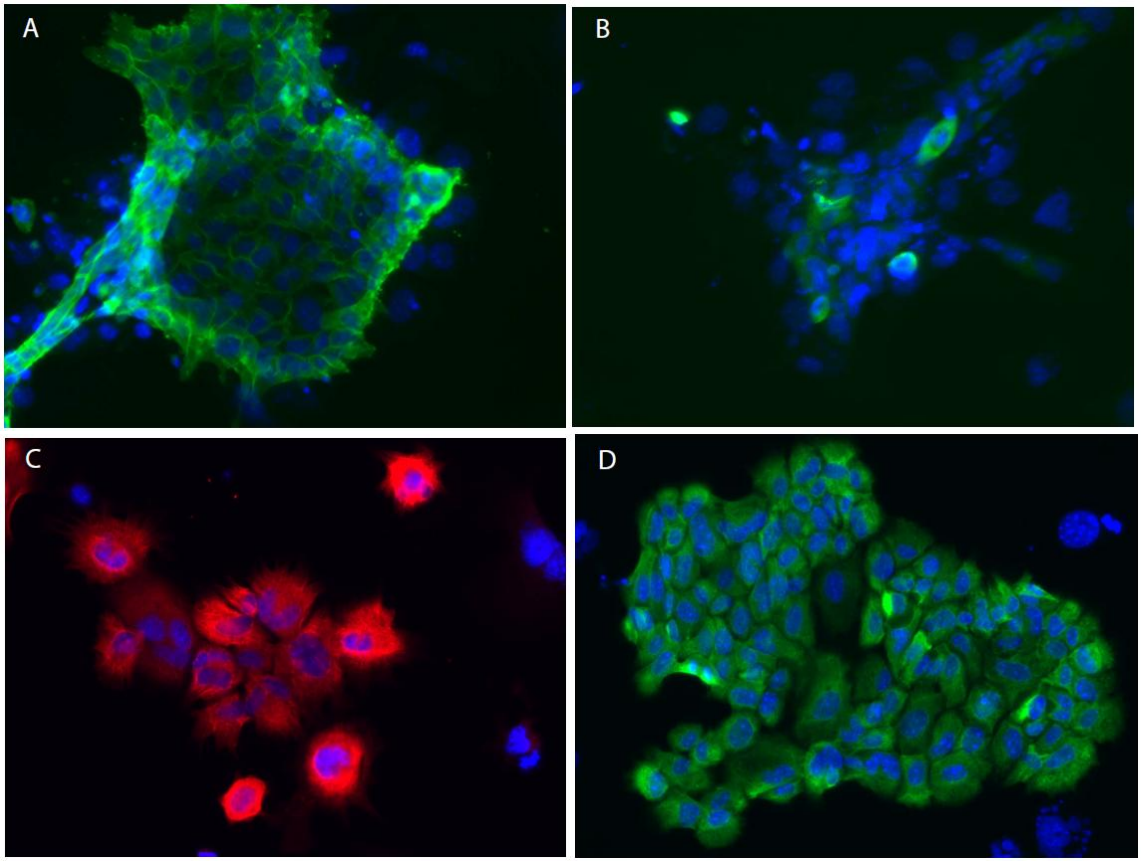



**Figure 3.6: *In vitro* colony immunofluorescence staining of ESA, KRT19, KRT14 and KRT8.**

A, ESA staining of luminal epithelial colony. B, KRT19 staining localizes to luminal epithelial colonies. C, KRT14 staining of a myoepithelial colony. D, KRT8 stains luminal epithelial colonies. In all images, DAPI blue staining demarcates nuclei. Internal negative controls may be observed by DAPI nuclei on periphery of colonies that did not stain positive for the indicated protein. All images takes at 100X magnification.

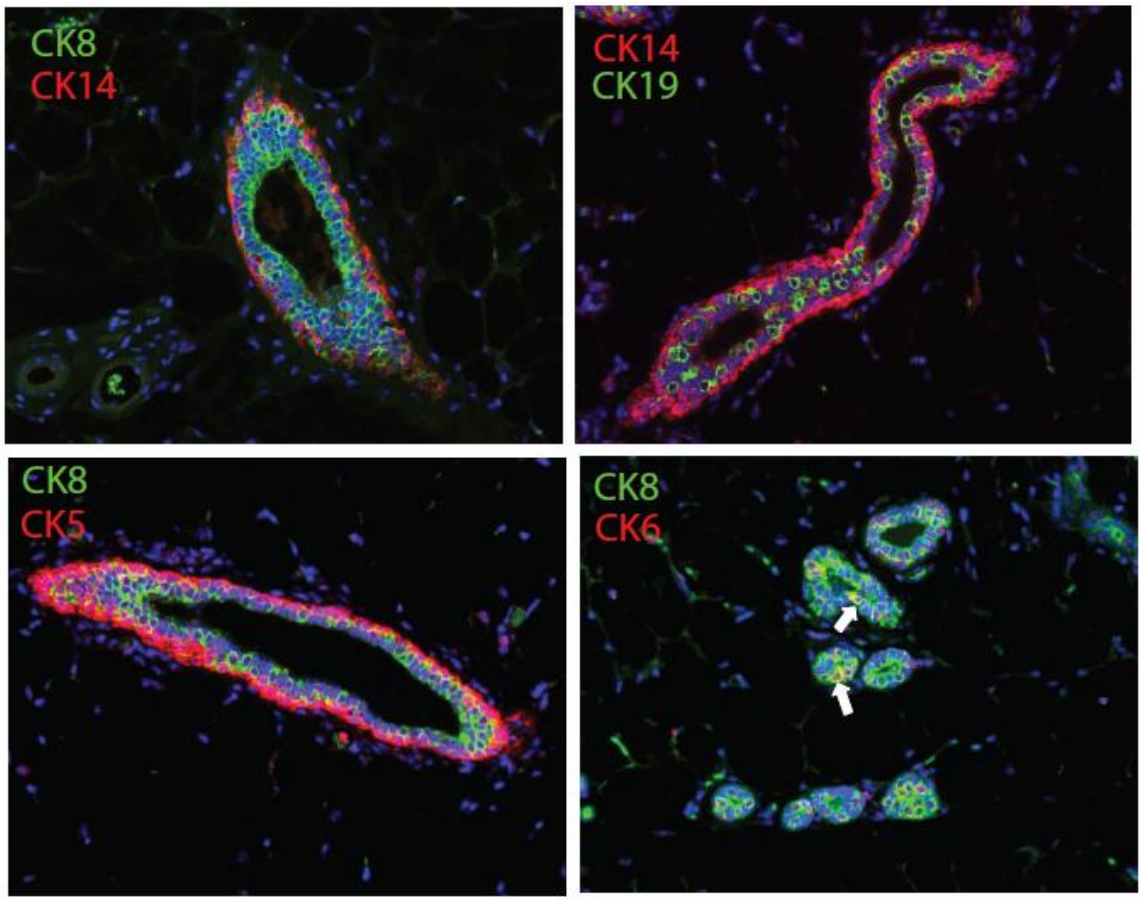



**Figure 3.7: Immunocytochemistry of virgin mammary tissue based on lineage specific keratin proteins.**

Paraffin sections of mammary fat pads were stained with primary antibodies against the indicated protein. Secondary antibodies were anti-rat Alexa 488 or anti-rabbit Alexa 594. DAPI blue staining was employed to distinguish nuclei. All images are 25X magnification.

| NAME | SIZE | ES | NES | NOM p-val | FDR q-val | FWER p-val |
|---|---|---|---|---|---|---|
| NEUROTRANSMITTER_BINDING | 49 | 0.53042716 | 2.126814 | 0 | 0.002048 | 0.007 |
| NEUROTRANSMITTER_RECEPTOR_ACTIVITY | 47 | 0.54636055 | 2.191138 | 0 | 0.002345 | 0.004 |
| ACETYLCHOLINE_BINDING | 15 | 0.6970162 | 1.976513 | 0 | 0.009354 | 0.073 |
| AMINE_BINDING | 21 | 0.6083901 | 1.989878 | 0 | 0.00959 | 0.06 |
| RHODOPSIN_LIKE_RECEPTOR_ACTIVITY | 105 | 0.43217456 | 1.999007 | 0 | 0.010969 | 0.052 |
| PEPTIDE_RECEPTOR_ACTIVITY | 49 | 0.4706282 | 1.892287 | 0 | 0.0177 | 0.212 |
| POTASSIUM_CHANNEL_ACTIVITY | 43 | 0.489236 | 1.899892 | 0.001548 | 0.018035 | 0.19 |
| G_PROTEIN_COUPLED_RECEPTOR_ACTIVITY | 149 | 0.38556328 | 1.900722 | 0 | 0.020945 | 0.19 |
| VOLTAGE_GATED_POTASSIUM_CHANNEL_ACTIVITY | 33 | 0.50742 | 1.842829 | 0.001704 | 0.027815 | 0.342 |
| ION_CHANNEL_ACTIVITY | 126 | 0.3817514 | 1.818775 | 0 | 0.031949 | 0.42 |
| GATED_CHANNEL_ACTIVITY | 105 | 0.37588215 | 1.774567 | 0 | 0.042155 | 0.583 |
| INOSITOL_OR_PHOSPHATIDYLINOSITOL_PHOSPHODIESTERASE_ACTIVITY | 12 | 0.6356808 | 1.781916 | 0.006369 | 0.042454 | 0.555 |
| SUBSTRATE_SPECIFIC_CHANNEL_ACTIVITY | 133 | 0.36812127 | 1.763776 | 0 | 0.044959 | 0.635 |

**Figure 3.8: GSEA analysis of molecular functions that differ between MASC and MYO cells.**

Analysis was performed using C57BL/6 microarrays on mammary epithelial populations. All arrays were RMA normalized and the data $\log_2$ transformed. The GSEA used the MSigDatabase file of all molecular functions as the a priori gene sets to determine which curated functional gene lists were different between the two populations. The samples included in the class list were 3 stem cell and 3 myoepithelial cell arrays. Of the resulting gene sets, only those that met the criteria of FDR<0.05 was used due to the small numbers of samples in each group of arrays. NAME, name of molecular pathway, SIZE, number of genes in list, ES, enrichment score, NES, normalized enrichment score, FDR q-value, the FDR statistic error rate, FWER, family-wise error rate statistic. FDR, which is less conservative than FWER, was used as the cutoff criteria since the GSEA algorithm focuses on hypotheses testing through permutation.

| | | | |
|---|---|---|---|
| Abca8a | Ezh1 | Rarres2 | 2610012C04Rik |
| Agtpbp1 | Gas7 | Rhpn2 | 4930515G01Rik |
| Agtr1a | Gas8 | Rpp14 | 4930515I15 |
| Ahnak | Gpm6b | S100a4 | 4930535C22Rik |
| Apool | Hod | Slc35f1 | 6530401C20Rik |
| Aspa | Hspa12a | Sypl2 | 9130014G24Rik |
| Atp1a2 | Inpp4b | Tcrg | A530047J11Rik |
| Atp1b2 | Irf8 | Tmem106a | A530088E08Rik |
| Atp2b4 | Kcna1 | Tmod2 | A830082K12Rik |
| BC030499 | Klf12 | Tnfaip2 | A930001M12Rik |
| Cadm2 | Lrrc27 | Trub1 | B230380D07Rik |
| Cadm4 | Mast1 | Ttyh1 | Cpe///LOC100046434 |
| Ccdc3 | Megf9 | Uhmk1 | ENSMUSG00000074335 |
| Ccl11 | Nat8l | Wnt8b | Gas2l3///LOC100047967 |
| Cog7 | Ngfr | Zdhhc2 | LOC100044153///OTTMUSG00000005491 |
| Cugbp2 | Nr2f2 | Zeb2 | LOC100046044///Nr2f1 |
| Dnpep | Pdzd2 | 0610042E11Rik | LOC666466///Obox2 |
| Edg3 | Prkcq | 0710005I19Rik | |

**Figure 3.9: Genes that are upregulated in MaSC versus MYO cells.**

Gene list was curated to remove probes that mapped to theoretical genes (i.e. NA) and multiple probes for the same gene were removed. This list represents a t-test using all unique permutations, standard Bonferroni correction. In addition, genes whose average expression did not differ by 2-fold or more between the two populations were removed post t-test. This process resulted in 71 unique genes.

| | | | | |
|---|---|---|---|---|
| Abca8a | Cyp4b1 | Lin7a | Tmem55a | D14Ertd668e |
| Ablim3 | Edg2 | Mal | Tmod2 | ENSMUSG00000074335 |
| Agps | Edg3 | Megf9 | Trub1 | Gp49a///Lilrb4 |
| Agtrl1 | Ednrb | Ngfr | Ttyh1 | LOC100041546///LOC100045833 ///Ly6c1///Ly6c2 |
| Ankrd49 | Exoc3l2 | Nr2f2 | Wdr89 | LOC100046044///Nr2f1 |
| Art3 | Flt1 | Pafah1b 1 | Zeb2 | LOC620695 |
| Asah3l | Gatm | Pdzd2 | 0610037M15Rik ///H2-Q6 | LOC627912///Mllt4///Zfp160 |
| Aspa | Gimap8 | Plp1 | 1110032E23Rik | LOC666466///Obox2 |
| Aspn | Gnb4 | Plxna2 | 2210419I08Rik | |
| Atp1a2 | Gpr63 | Pop1 | 2310043N10Rik | |
| Atp1b2 | Gucy1a3 | Prkcq | 2810022L02Rik | |
| C77583 | Hspa12a | Prrx1 | 2810436B12Rik | |
| Cadm2 | Idi2 | Rarres2 | 2810474O19Rik | |
| Ccl11 | Ifi47 | Rgs5 | 3110037L02Rik | |
| Chl1 | Ifnab | Sema7a | 4930549O18Rik | |
| Cldn5 | Inpp4b | Slc35f1 | 4933404K13Rik | |
| Clic5 | Irf8 | Snx10 | 4933406K04Rik | |
| Col27a1 | Kcna1 | Sparcl1 | 9130014G24Rik | |
| Cugbp2 | Klf12 | Thy1 | C130015C19 | |
| Cwf19l2 | Lama4 | Timp4 | C230081A13Rik | |

**Figure 3.10: Genes that are upregulated in MaSC versus MMPP cells.**

Gene list was curated to remove probes that mapped to theoretical genes (i.e. NA) and multiple probes for the same gene were removed. This list represents a permuted t-test with standard Bonferroni correction. In addition, genes whose average expression did not differ by 2-fold or more between the two populations were removed post t-test. This process resulted in 88 unique genes.

| Category | Term | P-Value | Fold Enrichment |
|---|---|---|---|
| GOTERM_CC_ALL | GO:0005886~plasma membrane | 0.030854 | 2.970009 |
| GOTERM_BP_ALL | GO:0007417~central nervous system development | 0.030044 | 10.28405 |
| GOTERM_BP_ALL | GO:0051480~cytosolic calcium ion homeostasis | 0.025346 | 73.41667 |
| GOTERM_BP_ALL | GO:0051179~localization | 0.007359 | 2.624628 |
| GOTERM_BP_ALL | GO:0051674~localization of cell | 0.006301 | 9.681319 |
| GOTERM_BP_ALL | GO:0006813~potassium ion transport | 0.014001 | 15.45614 |
| GOTERM_BP_ALL | GO:0022610~biological adhesion | 0.03342 | 5.189985 |
| GOTERM_BP_ALL | GO:0016337~cell-cell adhesion | 0.031349 | 10.04943 |
| GOTERM_BP_ALL | GO:0007155~cell adhesion | 0.03342 | 5.189985 |
| GOTERM_BP_ALL | GO:0006928~cell motility | 0.006301 | 9.681319 |
| GOTERM_BP_ALL | GO:0007610~behavior | 0.049476 | 7.819527 |
| GOTERM_BP_ALL | GO:0016477~cell migration | 0.040371 | 8.751656 |
| PANTHER_BP_ALL | BP00107:Cytokine and chemokine mediated signaling pathway | 0.005387 | 10.25481 |
| PANTHER_BP_ALL | BP00044:mRNA transcription regulation | 0.018111 | 1.776328 |
| PANTHER_BP_ALL | BP00103:Cell surface receptor mediated signal transduction | 0.045114 | 3.366854 |
| PANTHER_BP_ALL | BP00204:Cytokinesis | 0.045501 | 8.196078 |
| GOTERM_MF_ALL | GO:0016787~hydrolase activity | 0.026954 | 3.033152 |
| GOTERM_MF_ALL | GO:0022804~active transmembrane transporter activity | 0.002764 | 12.64633 |
| GOTERM_MF_ALL | GO:0030955~potassium ion binding | 0.003264 | 32.19594 |
| GOTERM_MF_ALL | GO:0031420~alkali metal ion binding | 0.008933 | 19.17681 |
| GOTERM_MF_ALL | GO:0017111~nucleoside-triphosphatase activity | 0.008252 | 8.554192 |
| GOTERM_MF_ALL | GO:0016887~ATPase activity | 0.002137 | 13.84362 |
| GOTERM_MF_ALL | GO:0043492~ATPase activity, coupled to movement of substances | 1.86E-04 | 32.04892 |
| GOTERM_MF_ALL | GO:0016820~hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances | 2.13E-04 | 30.58263 |
| GOTERM_MF_ALL | GO:0016818~hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides | 0.00951 | 8.123512 |

| | | | |
|---|---|---|---|
| GOTERM_MF_ALL | GO:0015075~ion transmembrane transporter activity | 0.016008 | 6.703643 |
| GOTERM_MF_ALL | GO:0016817~hydrolase activity, acting on acid anhydrides | 0.00983 | 8.025974 |
| GOTERM_MF_ALL | GO:0015405~P-P-bond-hydrolysis-driven transmembrane transporter activity | 2.44E-04 | 29.24464 |
| GOTERM_MF_ALL | GO:0015662~ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism | 0.001441 | 48.74107 |
| GOTERM_MF_ALL | GO:0015399~primary active transmembrane transporter activity | 2.44E-04 | 29.24464 |
| GOTERM_MF_ALL | GO:0016462~pyrophosphatase activity | 0.009285 | 8.194646 |
| GOTERM_MF_ALL | GO:0042626~ATPase activity, coupled to transmembrane movement of substances | 1.86E-04 | 32.04892 |
| GOTERM_MF_ALL | GO:0005215~transporter activity | 0.018969 | 4.263067 |
| GOTERM_MF_ALL | GO:0042625~ATPase activity, coupled to transmembrane movement of ions | 0.00309 | 33.10714 |
| GOTERM_MF_ALL | GO:0022892~substrate-specific transporter activity | 0.043077 | 4.587395 |
| GOTERM_MF_ALL | GO:0022891~substrate-specific transmembrane transporter activity | 0.023534 | 5.798194 |
| GOTERM_MF_ALL | GO:0042623~ATPase activity, coupled | 0.001311 | 16.41805 |
| GOTERM_MF_ALL | GO:0022857~transmembrane transporter activity | 0.004964 | 6.262236 |
| GOTERM_MF_ALL | GO:0005391~sodium:potassium-exchanging ATPase activity | 0.009487 | 194.9643 |
| PANTHER_MF_ALL | MF00230:Actin binding motor protein | 0.027807 | 5.576713 |
| PANTHER_MF_ALL | MF00031:Voltage-gated ion channel | 0.048012 | 7.955017 |
| PANTHER_MF_ALL | MF00042:Nucleic acid binding | 0.039681 | 1.831216 |
| PANTHER_MF_ALL | MF00024:Ion channel | 0.013779 | 15.60407 |

**Figure 3.11: Gene ontology of biological processes from DAVID database of MaSC specific genes.**

**BIBLIOGRAPHY**


1. Chang HY, Thomson JA, Chen X: **Microarray analysis of stem cells and differentiation**. *Methods Enzymol* 2006, **420**:225-254.
2. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, Wu L, Lindeman GJ, Visvader JE: **Generation of a functional mammary gland from a single stem cell**. *Nature* 2006, **439**:84-88.
3. Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, Li HI, Eaves CJ: **Purification and unique properties of mammary epithelial stem cells**. *Nature* 2006, **439**:993-997.
4. Brown PO, Botstein D: **Exploring the new world of the genome with DNA microarrays**. *Nat Genet* 1999, **21**:33-37.
5. Butte A: **The use and analysis of microarray data**. *Nat Rev Drug Discov* 2002, **1**:951-960.
6. Kuo WP, Jenssen TK, Butte AJ, Ohno-Machado L, Kohane IS: **Analysis of matched mRNA measurements from two different microarray technologies**. *Bioinformatics* 2002, **18**:405-412.
7. Langbein L, Rogers MA, Praetzel S, Aoki N, Winter H, Schweizer J: **A novel epithelial keratin, hK6irs1, is expressed differentially in all layers of the inner root sheath, including specialized huxley cells (Flugelzellen) of the human hair follicle**. *J Invest Dermatol* 2002, **118**:789-799.
8. Chung CH, Bernard PS, Perou CM: **Molecular portraits and the family tree of cancer**. *Nat Genet* 2002, **32 Suppl**:533-540.
9. Rhee SY, Wood V, Dolinski K, Draghici S: **Use and misuse of the gene ontology annotations**. *Nat Rev Genet* 2008, **9**:509-515.
10. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J, Thiagarajan M, White JA, Quackenbush J: **TM4 microarray software suite**. *Methods Enzymol* 2006, **411**:134-193.
11. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, et al.: **TM4: a free, open-source system for microarray data management and analysis**. *Biotechniques* 2003, **34**:374-378.
12. Asselin-Labat ML, Shackleton M, Stingl J, Vaillant F, Forrest NC, Eaves CJ, Visvader JE, Lindeman GJ: **Steroid hormone receptor status of mouse mammary stem cells**. *J Natl Cancer Inst* 2006, **98**:1011-1014.
13. Eblaghie MC, Song SJ, Kim JY, Akita K, Tickle C, Jung HS: **Interactions between FGF and Wnt signals and Tbx3 gene expression in mammary gland initiation in mouse embryos**. *J Anat* 2004, **205**:1-13.
14. Dudoit S, Gentleman RC, Quackenbush J: **Open source software for the analysis of microarray data**. *Biotechniques* 2003, **Suppl**:45-51.
15. Loganantharaj R, Cheepala S, Clifford J: **Metric for Measuring the Effectiveness of Clustering of DNA Microarray Expression**. *BMC Bioinformatics* 2006, **7 Suppl 2**:S5.
16. Rudland PS, Warburton MJ, Monaghan P, Ritter MA: **Thy-1 antigen on normal and neoplastic rat mammary tissues: changes in location and amount of antigen**

**during differentiation of cultured stem cells**. *J Natl Cancer Inst* 1982, **68**:799-811.
17. Bocker W, Bier B, Freytag G, Brommelkamp B, Jarasch ED, Edel G, Dockhorn-Dworniczak B, Schmid KW: **An immunohistochemical study of the breast using antibodies to basal and luminal keratins, alpha-smooth muscle actin, vimentin, collagen IV and laminin. Part I: Normal breast and benign proliferative lesions**. *Virchows Arch A Pathol Anat Histopathol* 1992, **421**:315-322.
18. Bartek J, Bartkova J, Taylor-Papadimitriou J: **Keratin 19 expression in the adult and developing human mammary gland**. *Histochem J* 1990, **22**:537-544.
19. Grimm SL, Bu W, Longley MA, Roop DR, Li Y, Rosen JM: **Keratin 6 is not essential for mammary gland development**. *Breast Cancer Res* 2006, **8**:R29.
20. Grimm SL, Rosen JM: **The role of C/EBPbeta in mammary gland development and breast cancer**. *J Mammary Gland Biol Neoplasia* 2003, **8**:191-204.
21. Welm B, Behbod F, Goodell MA, Rosen JM: **Isolation and characterization of functional mammary gland stem cells**. *Cell Prolif* 2003, **36 Suppl 1**:17-32.
22. Alvi AJ, Clayton H, Joshi C, Enver T, Ashworth A, Vivanco MM, Dale TC, Smalley MJ: **Functional and molecular characterisation of mammary side population cells**. *Breast Cancer Res* 2003, **5**:R1-8.
23. Smith GH, Mehrel T, Roop DR: **Differential keratin gene expression in developing, differentiating, preneoplastic, and neoplastic mouse mammary epithelium**. *Cell Growth Differ* 1990, **1**:161-170.
24. Sonnenberg A, Daams H, Van der Valk MA, Hilkens J, Hilgers J: **Development of mouse mammary gland: identification of stages in differentiation of luminal and myoepithelial cells using monoclonal antibodies and polyvalent antiserum against keratin**. *J Histochem Cytochem* 1986, **34**:1037-1046.
25. Moll R, Divo M, Langbein L: **The human keratins: biology and pathology**. *Histochem Cell Biol* 2008, **129**:705-733.
26. Zimek A, Weber K: **The organization of the keratin I and II gene clusters in placental mammals and marsupials show a striking similarity**. *Eur J Cell Biol* 2006, **85**:83-89.
27. Hesse M, Zimek A, Weber K, Magin TM: **Comprehensive analysis of keratin gene clusters in humans and rodents**. *Eur J Cell Biol* 2004, **83**:19-26.
28. Livasy CA, Karaca G, Nanda R, Tretiakova MS, Olopade OI, Moore DT, Perou CM: **Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma**. *Mod Pathol* 2006, **19**:264-271.
29. Vos JH, van den Ingh TS, Misdorp W, Molenbeek RF, van Mil FN, Rutteman GR, Ivanyi D, Ramaekers FC: **Immunohistochemistry with keratin, vimentin, desmin, and alpha-smooth muscle actin monoclonal antibodies in canine mammary gland: benign mammary tumours and duct ectasias**. *Vet Q* 1993, **15**:89-95.
30. Guelstein VI, Tchypysheva TA, Ermilova VD, Ljubimov AV: **Myoepithelial and basement membrane antigens in benign and malignant human breast tumors**. *Int J Cancer* 1993, **53**:269-277.

31. Troyanovsky SM, Guelstein VI, Tchipysheva TA, Krutovskikh VA, Bannikov GA: **Patterns of expression of keratin 17 in human epithelia: dependency on cell position**. *J Cell Sci* 1989, **93 ( Pt 3)**:419-426.
32. Dulbecco R, Allen WR, Bologna M, Bowman M: **Marker evolution during the development of the rat mammary gland: stem cells identified by markers and the role of myoepithelial cells**. *Cancer Res* 1986, **46**:2449-2456.
33. Yu H, Yuan Y, Shen H, Cheng T: **Hematopoietic stem cell exhaustion impacted by p18 INK4C and p21 Cip1/Waf1 in opposite manners**. *Blood* 2006, **107**:1200-1206.
34. Topley GI, Okuyama R, Gonzales JG, Conti C, Dotto GP: **p21(WAF1/Cip1) functions as a suppressor of malignant skin tumor formation and a determinant of keratinocyte stem-cell potential**. *Proc Natl Acad Sci U S A* 1999, **96**:9089-9094.
35. Li N, Zhang Y, Naylor MJ, Schatzmann F, Maurer F, Wintermantel T, Schuetz G, Mueller U, Streuli CH, Hynes NE: **Beta1 integrins regulate mammary gland proliferation and maintain the integrity of mammary alveoli**. *Embo J* 2005, **24**:1942-1953.
36. Platonova N, Scotti M, Babich P, Bertoli G, Mento E, Meneghini V, Egeo A, Zucchi I, Merlo GR: **TBX3, the gene mutated in ulnar-mammary syndrome, promotes growth of mammary epithelial cells via repression of p19ARF, independently of p53**. *Cell Tissue Res* 2007, **328**:301-316.
37. Bamshad M, Lin RC, Law DJ, Watkins WC, Krakowiak PA, Moore ME, Franceschini P, Lala R, Holmes LB, Gebuhr TC, et al.: **Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome**. *Nat Genet* 1997, **16**:311-315.
38. Davenport TG, Jerome-Majewska LA, Papaioannou VE: **Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in human ulnar mammary syndrome**. *Development* 2003, **130**:2263-2273.
39. Brummelkamp TR, Kortlever RM, Lingbeek M, Trettel F, MacDonald ME, van Lohuizen M, Bernards R: **TBX-3, the gene mutated in Ulnar-Mammary Syndrome, is a negative regulator of p19ARF and inhibits senescence**. *J Biol Chem* 2002, **277**:6567-6572.
40. Jacobs JJ, Keblusek P, Robanus-Maandag E, Kristel P, Lingbeek M, Nederlof PM, van Welsem T, van de Vijver MJ, Koh EY, Daley GQ, et al.: **Senescence bypass screen identifies TBX2, which represses Cdkn2a (p19(ARF)) and is amplified in a subset of human breast cancers**. *Nat Genet* 2000, **26**:291-299.
41. Lingbeek ME, Jacobs JJ, van Lohuizen M: **The T-box repressors TBX2 and TBX3 specifically regulate the tumor suppressor gene p14ARF via a variant T-site in the initiator**. *J Biol Chem* 2002, **277**:26120-26127.
42. Li Z, Tognon CE, Godinho FJ, Yasaitis L, Hock H, Herschkowitz JI, Lannon CL, Cho E, Kim SJ, Bronson RT, et al.: **ETV6-NTRK3 fusion oncogene initiates breast cancer from committed mammary progenitors via activation of AP1 complex**. *Cancer Cell* 2007, **12**:542-558.
43. Mine T, Matsueda S, Li Y, Tokumitsu H, Gao H, Danes C, Wong KK, Wang X, Ferrone S, Ioannides CG: **Breast cancer cells expressing stem cell markers**

**CD44(+) CD24 (lo) are eliminated by Numb-1 peptide-activated T cells**. *Cancer Immunol Immunother* 2008.
44. Lu ZQ, Li HG, Zhang HZ, Fan MJ, Shen XM, He XX: **[Expression and significance of CD44(+)ESA(+)CD24(-/low), stem cell markers for breast cancer, in non-small-cell lung carcinoma]**. *Ai Zheng* 2008, **27**:575-579.
45. Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, Wicha M, Clarke MF, Simeone DM: **Identification of pancreatic cancer stem cells**. *Cancer Res* 2007, **67**:1030-1037.
46. Gudjonsson T, Villadsen R, Nielsen HL, Ronnov-Jessen L, Bissell MJ, Petersen OW: **Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties**. *Genes Dev* 2002, **16**:693-706.
47. Stingl J, Eaves CJ, Kuusk U, Emerman JT: **Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast**. *Differentiation* 1998, **63**:201-213.
48. Dylla SJ, Beviglia L, Park IK, Chartier C, Raval J, Ngan L, Pickell K, Aguilar J, Lazetic S, Smith-Berdan S, et al.: **Colorectal cancer stem cells are enriched in xenogeneic tumors following chemotherapy**. *PLoS ONE* 2008, **3**:e2428.
49. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF: **Prospective identification of tumorigenic breast cancer cells**. *Proc Natl Acad Sci U S A* 2003, **100**:3983-3988.
50. Cho RW, Wang X, Diehn M, Shedden K, Chen GY, Sherlock G, Gurney A, Lewicki J, Clarke MF: **Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors**. *Stem Cells* 2008, **26**:364-371.
51. Dalerba P, Dylla SJ, Park IK, Liu R, Wang X, Cho RW, Hoey T, Gurney A, Huang EH, Simeone DM, et al.: **Phenotypic characterization of human colorectal cancer stem cells**. *Proc Natl Acad Sci U S A* 2007, **104**:10158-10163.
52. Farr A, Nelson A, Truex J, Hosier S: **Epithelial heterogeneity in the murine thymus: a cell surface glycoprotein expressed by subcapsular and medullary epithelium**. *J Histochem Cytochem* 1991, **39**:645-653.
53. Yang SJ, Ahn S, Park CS, Holmes KL, Westrup J, Chang CH, Kim MG: **The quantitative assessment of MHC II on thymic epithelium: implications in cortical thymocyte development**. *Int Immunol* 2006, **18**:729-739.
54. Ruther U, Wagner EF, Muller R: **Analysis of the differentiation-promoting potential of inducible c-fos genes introduced into embryonal carcinoma cells**. *Embo J* 1985, **4**:1775-1781.
55. Ichinose Y, Nozaki M, Morita T, Matsushiro A: **Identification of mouse endodermal cytoskeletal protein C (ENDOC) with cytokeratin No. 19**. *Biochem Biophys Res Commun* 1990, **167**:644-647.
56. Oshima RG, Howe WE, Klier FG, Adamson ED, Shevinsky LH: **Intermediate filament protein synthesis in preimplantation murine embryos**. *Dev Biol* 1983, **99**:447-455.
57. Saurat JH, Didierjean L: **The epidermal Merkel cell is an epithelial cell**. *Dermatologica* 1984, **169**:117-120.
58. Rosenstraus MJ, Sterman B, Carr A, Brand L: **Fibroblast feeder layers inhibit differentiation of retinoic acid-treated embryonal carcinoma cells by**

**increasing the probability of stem cell renewal**. *Exp Cell Res* 1984, **152**:378-389.
59. Bartolome MV, Casanova ML, Carricondo F, del Castillo E, Jorcano JL, Gil-Loyzaga P: **Abnormal cochlea linked to deafness in transgenic mice expressing human cytokeratin K8**. *Histol Histopathol* 2002, **17**:827-836.
60. Oshima RG: **Developmental expression of murine extra-embryonic endodermal cytoskeletal proteins**. *J Biol Chem* 1982, **257**:3414-3421.
61. Brulet P, Babinet C, Kemler R, Jacob F: **Monoclonal antibodies against trophectoderm-specific markers during mouse blastocyst formation**. *Proc Natl Acad Sci U S A* 1980, **77**:4113-4117.
62. Kemler R, Brulet P, Schnebelen MT, Gaillard J, Jacob F: **Reactivity of monoclonal antibodies against intermediate filament proteins during embryonic development**. *J Embryol Exp Morphol* 1981, **64**:45-60.
63. Garcia-Iglesias MJ, Perez-Martinez C, Bravo-Moral A, Ferreras-Estrada MC, Garcia-Fernandez R, Espinosa-Alvarez J, Escudero-Diez A: **Spontaneous mouse mammary tumours: incidence and cytokeratin expression**. *Res Vet Sci* 1997, **63**:85-89.
64. Wang A, Arantes S, Yan L, Kiguchi K, McArthur MJ, Sahin A, Thames HD, Aldaz CM, Macleod MC: **The transcription factor ATF3 acts as an oncogene in mouse mammary tumorigenesis**. *BMC Cancer* 2008, **8**:268.
65. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al.: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles**. *Proc Natl Acad Sci U S A* 2005, **102**:15545-15550.
66. Cohen S, Greenberg ME: **Communication between the synapse and the nucleus in neuronal development, plasticity, and disease**. *Annu Rev Cell Dev Biol* 2008, **24**:183-209.
67. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA: **DAVID: Database for Annotation, Visualization, and Integrated Discovery**. *Genome Biol* 2003, **4**:P3.
68. Ohta H, Aoyagi K, Fukaya M, Danjoh I, Ohta A, Isohata N, Saeki N, Taniguchi H, Sakamoto H, Shimoda T, et al.: **Cross talk between hedgehog and epithelial-mesenchymal transition pathways in gastric pit cells and in diffuse-type gastric cancers**. *Br J Cancer* 2009, **100**:389-398.
69. Korpal M, Lee ES, Hu G, Kang Y: **The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2**. *J Biol Chem* 2008, **283**:14910-14914.
70. Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB, Pena R, Bonilla F, de Herreros AG: **A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition**. *Genes Dev* 2008, **22**:756-769.
71. Kurihara I, Lee DK, Petit FG, Jeong J, Lee K, Lydon JP, DeMayo FJ, Tsai MJ, Tsai SY: **COUP-TFII mediates progesterone regulation of uterine implantation by controlling ER activity**. *PLoS Genet* 2007, **3**:e102.

72. Saitoh T, Mine T, Katoh M: **Expression and regulation of WNT8A and WNT8B mRNAs in human tumor cell lines: up-regulation of WNT8B mRNA by beta-estradiol in MCF-7 cells, and down-regulation of WNT8A and WNT8B mRNAs by retinoic acid in NT2 cells**. *Int J Oncol* 2002, **20**:999-1003.
73. Katoh M: **Conserved POU/OCT- and GATA-binding sites in 5'-flanking promoter region of mammalian WNT8B orthologs**. *Int J Oncol* 2007, **30**:1273-1277.
74. Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, Hartley L, Robb L, Grosveld FG, van der Wees J, et al.: **Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation**. *Nat Cell Biol* 2007, **9**:201-209.
75. Benhaj K, Akcali KC, Ozturk M: **Redundant expression of canonical Wnt ligands in human breast cancer cell lines**. *Oncol Rep* 2006, **15**:701-707.
76. Parsa S, Ramasamy SK, De Langhe S, Gupte VV, Haigh JJ, Medina D, Bellusci S: **Terminal end bud maintenance in mammary gland is dependent upon FGFR2b signaling**. *Dev Biol* 2008, **317**:121-131.
77. Stingl J, Emerman JT, Eaves CJ: **Enzymatic dissociation and culture of normal human mammary tissue to detect progenitor activity**. *Methods Mol Biol* 2005, **290**:249-263.
78. Srinivasan K, Strickland P, Valdes A, Shin GC, Hinck L: **Netrin-1/neogenin interaction stabilizes multipotent progenitor cap cells during mammary gland morphogenesis**. *Dev Cell* 2003, **4**:371-382.
79. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S: **AmiGO: online access to ontology and annotation data**. *Bioinformatics* 2009, **25**:288-289.
80. **The Gene Ontology project in 2008**. *Nucleic Acids Res* 2008, **36**:D440-444.
81. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, et al.: **PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification**. *Nucleic Acids Res* 2003, **31**:334-341.
82. Stingl J, Eaves CJ, Zandieh I, Emerman JT: **Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue**. *Breast Cancer Res Treat* 2001, **67**:93-109.
83. Hermosilla T, Munoz D, Herrera-Molina R, Valdivia A, Munoz N, Nham SU, Schneider P, Burridge K, Quest AF, Leyton L: **Direct Thy-1/alphaVbeta3 integrin interaction mediates neuron to astrocyte communication**. *Biochim Biophys Acta* 2008, **1783**:1111-1120.
84. Avalos AM, Labra CV, Quest AF, Leyton L: **Signaling triggered by Thy-1 interaction with beta 3 integrin on astrocytes is an essential step towards unraveling neuronal Thy-1 function**. *Biol Res* 2002, **35**:231-238.
85. Leyton L, Schneider P, Labra CV, Ruegg C, Hetz CA, Quest AF, Bron C: **Thy-1 binds to integrin beta(3) on astrocytes and triggers formation of focal contact sites**. *Curr Biol* 2001, **11**:1028-1038.
86. Dreyer EB, Leifer D, Heng JE, McConnell JE, Gorla M, Levin LA, Barnstable CJ, Lipton SA: **An astrocytic binding site for neuronal Thy-1 and its effect on neurite outgrowth**. *Proc Natl Acad Sci U S A* 1995, **92**:11195-11199.

87. Cray C, Keane RW, Malek TR, Levy RB: **Regulation and selective expression of Ly-6A/E, a lymphocyte activation molecule, in the central nervous system**. *Brain Res Mol Brain Res* 1990, **8**:9-15.
88. Mizobe T, Kidoguchi K, Tamaki M, Sasayama T, Kondoh T, Kohmura E: **Sca-1 and Thy-1 accelerate neuron-like differentiation in bone marrow stromal cells**. *Kobe J Med Sci* 2006, **52**:119-125.

# CHAPTER 4

# EFFECTS OF ROS AND TGFB ON MAMMARY PROGENITORS

## SUMMARY

The murine mammary system is a complex system comprised of multiple cell types. There exists a cellular hierarchy in the system, similar to many other somatic tissues. We have recently described the murine mammary stem cells by the surface protein phenotype. In the studies presented here, we explore mechanisms of stem cell protection and differentiation. Our data shows that a stem cell enriched population of mammary cells has lower ROS levels than further differentiated progeny. In addition, stem cells with the lowest amount of ROS have a significant growth advantage for *in vivo* duct regeneration. We also performed studies on TGFβ signaling in isolated mammary populations. There has been significant evidence implicating TGFβ signaling as a major regulator of mammary development. In addition, corrupted TGFβ signaling is often found in human breast cancers. To assess the role of TGFβ signaling in mammary stem cells, we cultured various mammary populations in the presence of human TGFβ1 ligand. Our results show TGFβ1 ligand stimulates a motility program in early MaCFC progenitor cells. Intriguingly, TGFβ1 seemed to have no effect on MaSC and MMPP cells *in vitro*.

## INTRODUCTION

In this chapter, we explore some mechanisms of mammary stem cell protection and differentiation. Our previous work defined a new phenotype for murine mammary stem cells. The isolation of a mammary stem cell population allowed us to compare their functionality with further differentiated progeny. This reasoning was applied to two distinct hypotheses. The first was stem cells are long lived so they must protect themselves from potential sources of genotoxins such as reactive oxygen species. The second was TGFβ signaling affects mammary stem cells differently than other populations.

### ROS in mammary stem cells

Reactive Oxygen Species (ROS) is a general term that applies to free radicals (molecules with an unpaired electron) and a variety of other molecules that are all derived from molecular oxygen[1]. Oxygen in its ground state is in the triplet state, meaning it is a bi-radical that contains two unpaired electrons in the outer orbital shell. Since both of the unpaired electrons have the same spin, oxygen can only react with one electron at a time and thus is not very reactive. However, if one of the unpaired electrons becomes excited and reverses or changes its spin, the resulting oxygen molecule becomes an oxidant since the two electrons can now react with other pairs of electrons.

Understanding the reduction process of oxygen is paramount to understanding oxidative stress. When oxygen is reduced by one electron, a superoxide anion intermediate is formed, the major precursor of most ROS. This intermediate also plays a regulatory role in oxidative chain reactions. Further dismutation of the superoxide anion either spontaneously or via superoxide dismutase produces hydrogen peroxide which, in turn, can be completely reduced to water. Hydrogen peroxide can also be partially reduced to hydroxyl radical, which is known to be one of the most powerful naturally occurring oxidants. Superoxide anions can also react with other radicals including, but not limited to, nitric oxide producing peroxynitrite, another strong oxidant. These types of molecular have been named reactive nitrogen species (RNS).

Oxidative stress is a term that describes a situation where there is an excess of ROS and/or RNS and a limitation of antioxidant defenses[2]. Interestingly, small changes in ROS may facilitate intracellular signaling. Large increases in intracellular ROS lead to modification of proteins, polysaccharides and DNA. Superoxide anions are found in many cell types and may be produced both enzymatically and non-enzymatically[3]. One of the largest sources of ROS in tissues comes from the mitochondrial electron transport chain in which numerous redox centers may leak electrons to oxygen. Superoxide formation localizes to the outer mitochondrial membrane, in the matrix and on both sides of the inner mitochondrial membrane. ROS and RNS may be detected by a number of techniques including fluorometry, chemiluminescence and electron paramagnetic resonance[4]. Unfortunately, these methods are susceptible to artifacts since the detection is not specific for a particular species of ROS or RNS, and the probe may itself have reactive intermediates.

Mitochondria have various defense mechanisms against ROS[5,6]. For example, superoxide anions can be converted to hydrogen peroxide by the metalloenzyme family of the superoxide dismutases (SOD). Both SOD and cytochrome c participate is reducing intra-mitochondrial ROS levels[7]. In addition, enzymes involved with glutathione peroxidase activity reduce ROS levels, specifically in liver cells. Catalase, an enzyme that reduces hydrogen peroxide in peroxisomes, is also a major ROS regulator in heart cells and skeletal muscle[8-10]. Oxidative stress can result not only from transient excess of ROS but also deficiencies in antioxidant defense mechanism via genetic predisposition or aging.

Since 1956, when Harman proposed the "free radical theory of aging" essentially saying that ROS determines the lifetime of an organism, scientists have been investigating how ROS affects cells[11]. Given that stem cells are long lived components of many somatic tissues and are responsible for their long-term maintenance, their effective processing of ROS may determine how successful they perform their normal physiologic functions[12,13]. Steady-state increases of ROS in cells over long periods of time have been shown to have deleterious effects[14-16]. In the blood system, the best studied of all somatic tissues, the HSCs protection from ROS is crucial for their self-renewal capacity[17,18]. These effects could be reversed using antioxidant therapy, showing the specificity of ROS damage in the HSC compartment. As HSCs differentiate, the protective mechanisms that are so critical for HSC maintenance are lost, presumably to begin generation of higher endogenous ROS as a bactericidal mechanism[19].

**TGF Beta Signaling**

The transforming growth factor B (TGFβ) superfamily is an evolutionarily conserved set of secreted peptides that participate in the regulation of almost all aspects of cellular behavior[20,21]. There are numerous members of the family including three isoforms of TGFβ (TGFβ1. TGFβ2, TGFβ3), bone morphogenic factors (BMPs), growth and differentiation factors (GDFs), activins, inhibins, and nodal. The TGFβ signaling pathway is shown in Figure 4.1. The secreted ligand TGFβ is produced as a prodomain containing protein. It is secreted as an inactive 25kDa mature peptide in a non-covalent association with the N-terminal prodomain of the precursor protein. The TGFβ ligand signals through a receptor complex composed of type I (TGFβRI) and type II (TGFβRII) serine threonine kinases. Typically, TGFβ binds to a TGFβ type II receptor dimer, which in turn catalyzes the phosphorylation of the Type I receptor in a hetero-tetrameric complex. In mammals, there are seven type II receptors and five type I receptors, adding complexity to the system. The BMPs bind to the cognate receptor BMP receptor 2 (BMPR2)[22]. The downstream effectors of the TGFβ ligands are the SMAD family of transcription factors.

The TGFβ superfamily is critical to mammary gland development[23-26]. Multiple members have been found in both ducts and lobules of the mammary epithelium. Interestingly, the levels of TGFβ decrease significantly during lactation as opposed to earlier times in pregnancy. In fact, the ligands are differentially expressed immediately after lactation through progression to involution. Type I and II receptors were localized to both epithelium and stroma during mammary development. TGFβ1 or TGFβ3 has growth suppressive effects on lobules found at the leading edge of the growing ductal system. Further confirmation of this observation was found in *MMTV*-

*TGFβ1* transgenic mice where the mammary gland was hypoplastic, there was delayed invasion of the ductal epithelium into the stromal fat pad, and there was an absence of side branching[27]. In addition, the growth rates of the mammary tree were 2-4 fold reduced compared to controls depending on the age of the mice examined. This result paralleled results of *Smad3*$^{-/-}$ mammary development[28]. The converse experiment confirmed the role of TGFβ in ductal morphogenesis. In *TGFβ1*$^{+/-}$ heterozygous mice, increased proliferation of both lobules and ducts was observed, as well as alveolar development[29]. A dominant negative active TGFβRII mouse also demonstrated increased ductal proliferation, suggesting the effects of TGFB1 ligand functions through TGFβRII signaling and downstream Smad3 activation. *WAP-Tgfb1* mice, in which TGFβ1 should be turned on during pregnancy, had defective secretory alveolar differentiation resulting in early apoptosis[30-32]. *MMTV-TGFβRI* mice also had increased apoptosis in alveolar cell containing lobules[27,30,33,34]. Interestingly, activating TGFβRII signaling under the control of the MMTV promoter initiated early secretory alveolar development in virgin mice. It was suggested that this effect may have been due to the accumulation of alveolar cells since the phenotype was only observed after weeks of age.

Using a transgenic mouse model driving *TGFβ3* under control of the B-lactoglobulin promoter was found to increase apoptosis in alveolar cells through a Smad4 signaling pathway and concomitant phosphorylation of Stat3[35]. *TGFβ3*$^{-/-}$ mammary tissue had a decrease in alveolar apoptosis. Dominant negative TGFβRII in mammary epithelium (*MMTV-DNIIR*) resulted in delayed involution by a failure to phosphorylate Akt and Forkhead transcription factor 1 (*FKHR*)[36,37]. Interestingly, TGFβ3 production

was localized to myoepithelial cells during involution, corresponding to sites of increased laminin, a molecule found at the basal lamina dividing the myoepithelial cells from the stroma. *WAP-Tgfβ1* mice had a failure of epithelial cells to grow a mammary ductal system. TGFβ1 inhibited the self-renewal of PI-MEC. Smad4 levels increase during proliferation and alveolar differentiation. Work performed in tumor models suggested that TGFβ is involved in the differentiation and transformation of myoepithelial cells.

## MATERIALS AND METHODS

### Animals

All animals used in the study were C57Bl/6 or 129S1/SvImJ mice that were maintained at the Stanford Animal Facility in accordance with the guidelines of both Institutional Animal Care Use Committees.

### Murine mammary stem cell isolation

Mammary glands from 6–12-week-old female C57BL/6J or 129S1/SvImJ mice were dissociated as described[10] with minor modifications. Specifically, mammary fat pads were collected and placed directly into Medium 199 (Gibco BRL) supplemented with 20 mM HEPES and penicillin, streptomycin and actinomycin. Tissue was minced using sterile razor blades and 4 Wünsch units of Liberase Blendzyme 4 (Roche) and 100 Kunitz units of DNase I (Sigma) were added. Tissue was incubated for 60–90 min in a 37 °C and 5% $CO_2$ incubator, during which the cells were mechanically aspirated every 30 min. Cells were pelleted by centrifugation for 5 min at 4 °C and 350*g*. After lysis of

the red blood cells with ACK lysis buffer (Gibco BRL), a single-cell suspension was obtained by further enzymatic digestion for 2 min in 0.25% trypsin, followed by another ~2 min in 5 mg/ml dispase II (StemCell Technologies) plus 200 Kunitz units DNase I (Sigma). Cells were then filtered through 40-um nylon mesh, pelleted and resuspended in staining media (HBSS and 2% heat-inactivated calf serum (HICS)). Cells were counted using trypan blue dye exclusion.

**Cell staining and flow cytometry**

Cells were stained at a concentration of $1 \times 10^6$ cells per 100 ul of HBSS with 2% HICS (staining media). Cells were blocked with rabbit or mouse IgG (1 mg/ml) at 1:100 dilution and antibodies were added at appropriate dilutions determined from titration experiments. For the normal mammary stem cell experiments, antibodies included CD49f, CD31, CD45, Ter119 (BD Pharmingen), CD24, Thy1.2 and CD140α (eBioscience). Cells were stained for 20 min on ice and washed with staining media. Cells were further stained with streptavidin-conjugated fluorophores and washed. Ultimately, cells were resuspended in staining media containing 7-aminoactinomycin D (1 µg/ml final concentration) or 4'-6-diamidino-2-phenylindole (DAPI, 1 µg/ml final concentration) to stain dead cells.

For all experiments, cells were analyzed and sorted using a FACSAriaII cell sorter (BD Bioscience). Side scatter and forward scatter profiles were used to eliminate debris and cell doublets. Dead cells were eliminated by excluding DAPI positive cells,

whereas contaminating human or mouse Lineage$^+$ cells were eliminated by excluding cells labeled with the fluorophore used for the lineage antibody cocktail. In cell-sorting experiments, cell populations underwent two consecutive rounds of purification (double sorting) when the initial purity was not deemed high enough and a sufficient number of cells were available. Final purities ranged from ~60% to >95%.

**Normal mammary stem cell DCF-DA transplant experiments**

CD24$^{med}$ CD49f$^{high}$ Lin$^-$ mammary cells (enriched for mammary repopulating units) were isolated from mammary fat pads from C57BL/6J female mice as described above. Cells were loaded with 10 uM DCF-DA (Invitrogen), incubated at 37 °C for 30 min, and sorted into 'ROS low' and 'ROS mid' sub-populations on the basis of their DCF-DA staining profile (in comparison to that of CD24$^{high}$ CD49f$^{low}$ Lin$^-$ progenitor cells, which showed an 'ROS high' profile). Mammary glands of 21-day-old female C57BL/6J mice were cleared of endogenous epithelium as previously described[10], and sorted cells were injected into each cleared fat pad using a Hamilton syringe. Injected glands were removed for wholemount analysis after 5–6 weeks. Transplants were scored as positive if epithelial structures consisting of ducts with lobules and/or terminal end buds and that arose from a central point were present.

***In vitro* 3-D cultures**

Cell populations were isolated as previously described. Freshly sorted cells were resuspended with Matrigel (BD) and allowed to gellate at 37 °C for 30 min. Epicult (Stem Cell Technologies) supplemented with 5%FBS and PSA was added. Human recombinant TGFβ1 ligand (R&D Systems) was added at a concentration of 8 ng/ml to cultures. Media was changed every two days. Experiment was allowed to proceed for 10 days and then colonies were counted and imaged. For GFP pictures, donor cells were obtained from pCx-GFP mice bred in-house.

## RESULTS

### Mammary stem cells with low ROS have a growth advantage *in vivo*

We began by asking whether low ROS concentrations that appear to be critical to self renewal of hematopoietic stem cells (HSCs)[19,38] are also a property of mammary epithelial stem cells[39,40]. Investigation of our microarray data revealed MaSCs, MMPPs and MaCFCs differentially expressed a number of genes involved in ROS (Figure 4.2). The MaSC and MaCFC arrays were the most dissimilar, and therefore we predicted they may have differences in their intracellular ROS levels. We chose not to use Thy-1 as a marker of stem cells for these studies as we only wanted to assess progenitors that engrafted *in vivo* versus those that could not. Since intracellular ROS may be detected using dyes that react when they come in contact with free radicals, we isolated $CD24^{med}CD49f^{high}Lin^{-}$ (MRU) mammary cells and $CD24^{high}CD49f^{low}Lin^{-}$ (MaCFC) early progenitor cells by flow cytometry. We measured intracellular concentrations of pro-oxidants using 2'-7'-dichlorofluorescein diacetate (DCF-DA) staining[38]. Sorted cells were immediately incubated with the DCF-DA stain for 30

minutes and then re-analyzed by flow cytometry. We went on to test ROS levels of MRU-enriched cells in two strains of mice. C57Bl/6 mice had lower ROS levels in the MRU-enriched population as compared to further differentiated MaCFC progenitor cells (Figure 4.3A). This result was similar to our findings in sorted populations from 129SS1/SvImJ mice (Figure 4.3B). The analysis of all experiments showed that CD24medCD49fhi cells did have lower average ROS expression (Figure 4.2C). These results were similar to those in HSCs, where lower ROS levels are thought to be protective[19,41,42].

Specifically, the MRU-enriched populations displayed low to intermediate ROS levels, while the progenitor-enriched populations contained more uniformly high levels of ROS. Similarly, analysis of the two populations with MitoSOX Red, a highly selective detection method for mitochondrial superoxide, revealed lower superoxide levels in the MRU-enriched population[12] (data not shown). Although MRUs had lower ROS than the MaCFCs, we noticed there were $ROS^{hi}$ and $ROS^{low}$ populations within the MRU fraction. The data suggested $ROS^{low}$ MRUs have greater protective mechanisms activated to reduce intracellular ROS. We hypothesized that MRUs with lower intracellular ROS may also have a growth advantage *in vivo* compared to MRUs with higher ROS. We did not transplant the MaCFC population since it is relatively devoid of engraftment potential[40].

We transplanted $CD24^{med}CD49f^{high}Lin^{-}$ cells based on their levels of DCF-DA staining. Cells were first sorted based on surface phenotype as described and then stained with DCF-DA. Cells were then re-sorted only based on their ROS profile since surface protein expression was dramatically altered during the DCF-DA staining procedure.

Interestingly, we found mammary stem cells with both low and intermediate ROS levels gave rise to epithelial outgrowths when transplanted into cleared fat pads (Figure 4.4). Similar heterogeneity of ROS concentrations was recently demonstrated in HSC-enriched populations[19,43], where it may have functional significance in modulating the HSC-niche interaction[44]. Taken together, our results demonstrate the first report of a protective mechanism of mammary stem cells based on the surface marker phenotype isolation. In addition, reduced ROS levels may be a shared defense mechanism in somatic stem cells.

**Ectopic TGFβ1 ligand has an effect on a specific subset of mammary epithelial cells**

Recently, a number of studies have shown TGFβ regulates the extracellular matrix by generation of ROS and calcium influx[45,46]. We probed our microarray data for TGFβ ligands and receptors (Figure 4.5) to investigate how signaling may be regulated differently in mammary populations. This analysis suggested the MaSC population downregulated some members, such as TgfBR1, but upregulated the downstream effector protein Tgfbi. These results suggested TGFβ signaling may have different effects on individual mammary populations.

To assess the role of TGFβ signaling in mammary epithelial populations, we added recombinant human TGFβ1 ligand to 3-D matrigel cultures of unsorted mammary epithelial cells. We used a high concentration of ligand as we hoped to elicit the most dramatic effect in culture, and to overcome the diffusion gradient inherent in adding soluble protein to the culture media and not embedded in the matrix. To visualize our results and to accurately count the resulting colonies, we used pCx-GFP donor cells.

These mice express GFP ubiquitously in mammary epithelium[47], enabling easy colony counting and morphological analysis. Due to the three dimensional nature of matrigel cultures, accurate colony formation assessment is challenging[48]. To investigate any potential effect of TGFβ1, we cultured 150K bulk unsorted epithelial cells with and without ligand. In wild-type cultures, we observed excellent colony formation, which consisted of either hollow or solid colonies (Figure 4.6, Top). When TGFβ was added to the culture media, we observed a dramatic decrease in colony formation (Figure 4.6, Bottom), and the induction of spindle-like myoepithelial differentiation. Some colonies were completely unaffected by ligand addition however, but others displayed more mesenchymal transition. We did observe mesenchymal looking cells in both hollow and solid colonies. These results suggested that TGFβ1 ligand had an effect on particular cell types within the culture, and thus fractionating the cells based on surface phenotype may prove useful to determine which cell type were responsive.

We next isolated MaSC (Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$) cells and plated with and without TGFβ1 to investigate how colony formation would be affected. We did not observe any significant difference of stem cell colonies when ligand was added compared to wild-type cultures (Figure 4.7), suggesting the stem cells are either irresponsive to this ligand or they have an autocrine TGFβ signaling mechanism that cannot be exogenously over-activated. The same was true for mammary multipotent progenitors in cell culture (Figure 4.7). These results suggest TGFβ1 ligand signaling does not directly affect differentiation of stem and MMPP cells.

We also isolated 5K or 10K MaCFC cells, plated them into Matrigel culture and added TGFβ1 ligand to investigate the effects on further differentiated bipotent

progenitors. In wild type cultures, we observed robust colony formation consistent with previously published reports[39]. Colonies were generally large and hollow, much bigger than colonies derived from either stem or MMPPs (Figure 4.8, Top). The addition of TGFβ1 ligand caused a significant reduction in the number of colonies produced by similar amounts of cells, and this effect was independent of the number of cells plated (Figure 4.8, Bottom). Colonies that did grow were much smaller than wild-type colonies, approximate in size to standard stem cell colonies. Upon closer inspection of these colonies, we routinely observed motile cells from the point of initial colony formation. These cells did not appear to be separate cells that had grown close to the initial colony, and the effect was noted for all colonies in the culture. Also, the motile cells did not display a mesenchymal phenotype, as would be expected. We also did not observe myoepithelial differentiation in our cultures, only migration. This ligand may not directly affect MaCFCs, since this assay does not show that every cell within the affected colonies retained their bipotentiality or surface marker phenotype. It is possible that MaCFCs differentiate in these culture conditions, and the migratory cells are a differentiated population. Taken together, these data demonstrated TGFβ1 ligand induced a motility program in some MaCFC-derived cells.

Since TGFβ1 had an effect on MaCFC colonies, we went on to test the responsiveness of differentiated luminal epithelial cells (EPI) to make colonies when ligand was added. In wild type cultures, we observed excellent colony formation when 5K and 10K cell were plated. This is the first report of plating isolated differentiated luminal cells in the Matrigel culture system. Resulting colonies looked similar to MaCFC colonies both in shape and size, but over time the colonies began to bulge (Figure 4.9,

Top). Colonies typically grew to a certain size as a spheroid and then began to expand in a different way. When TGFβ1 ligand was added, we observed a decrease of colony formation at 5K cells level and a significant decrease at 10K cells (Figure 4.9, Bottom). The effect of TGFβ1 addition on EPI cells was different than in MaCFC cultures. Colonies grew as wild type, except the second cell proliferation stage was completely abolished, retaining the boundaries of the initial spheroid formed (Figure 4.9, Top).

## DISCUSSION

Based on our recent definition of the murine mammary stem cell, or MaSC, we embarked on two studies to investigate functions of primitive cells. In our first study, we showed for the first time that mammary stem cells have a much lower amount of intracellular ROS than more differentiated cells. Also, stem cells with the lowest ROS had the best *in vivo* engraftment. Although we cannot rule out that chemical intermediates of the DCF-DA stain could have altered the results of our analysis, our results are similar to those found in HSCs in the blood[44,49], suggesting ROS tolerance may be a shared mechanism for stem cell protection. Further investigation remains to elucidate the genes responsible for the mammary stem cells' resistance to ROS. This population could upregulate enzymes that process the ROS faster or efflux pumps that get rid of the ROS molecules before they can do the cells damage. Our microarray data showed Abca8a, a new member of the ATP binding cassette transporter family, was upregulated only in the mammary stem cells enriched population. This gene was mainly expressed in the liver and the heart, two organs with well studied ROS models. In fact, Abc8a was one of the

only transporters that were not susceptible to damage by acute digoxin[50], suggesting it has a functional role in digoxin (and subsequently ROS) metabolism. We speculate reduction of Abca8a in the stem cell compartment may sensitize the stem cells to ROS damage.

In the second study we tried to determine if TGFβ signaling has different effects on the various mammary cell types using *in vitro* 3-D culturing. Our data revealed TGFβ1 did not affect the stem or multipotent cells in culture, but had dramatic effects on the MaCFC and EPI cells. From our results, it is unclear if this effect was limited to MaCFCs and EPI because of their luminal identity, as our previous gene expression results demonstrated. TGFβ signaling is commonly over-activated in luminal A human breast tumors[51]. Induction of the motility program in the MaCFCs suggests these cells may have initiated a mesenchymal transcriptional program. In a recent report from our group, we described a cancer stem cell population within *MMTV-Wnt1* mammary tumors[52]. This report described the cancer stem cell population as having a basal phenotype, and the non-cancerous cells within the same tumor had a more luminal phenotype. Therefore, the there may be a requirement that tumor initiating cells in the breast tumors initiate a mesenchymal phenotype during initial transformation. Alternatively, tumor initiating cells may begin with a luminal phenotype and then undergo a transition to a mesenchymal phenotype to maintain and expand tumorigenicity. In *MMTV-c-rel* transgenic mice, spontaneous breast tumors arise after a long latency period. A cell line created from these tumors displayed characteristics of extremely aggressive tumors, EMT and increased motility when subjected to the carcinogen DMBA, suggesting tumor phenotype directly correlates with presence of mesenchymal

cells[53]. However, more work remains to better characterize our MaCFC results, including *in vivo* transplantation of these cells. An interesting approach may be to infect these cells with an oncogene (such as *Ras*) to determine if TGFβ1 induces a true epithelial-to-mesenchymal transition. Also, measuring the intracellular ROS levels of mammary stem cells that have been subjected to ectopic TGFβ ligand may provide a molecular mechanism of genotoxic protection in these cells.

## FIGURES

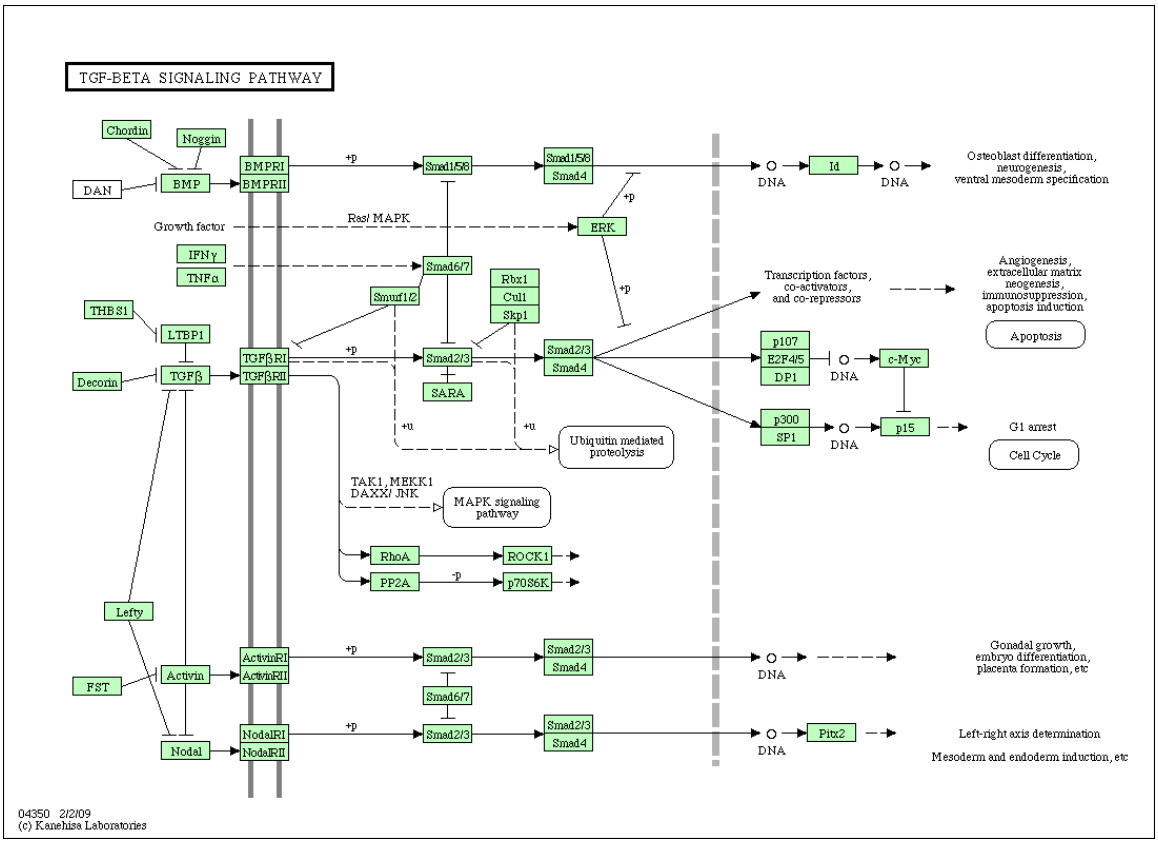



**Figure 4.1: The TGFβ Signaling Pathway according the KEGG Ontology.**

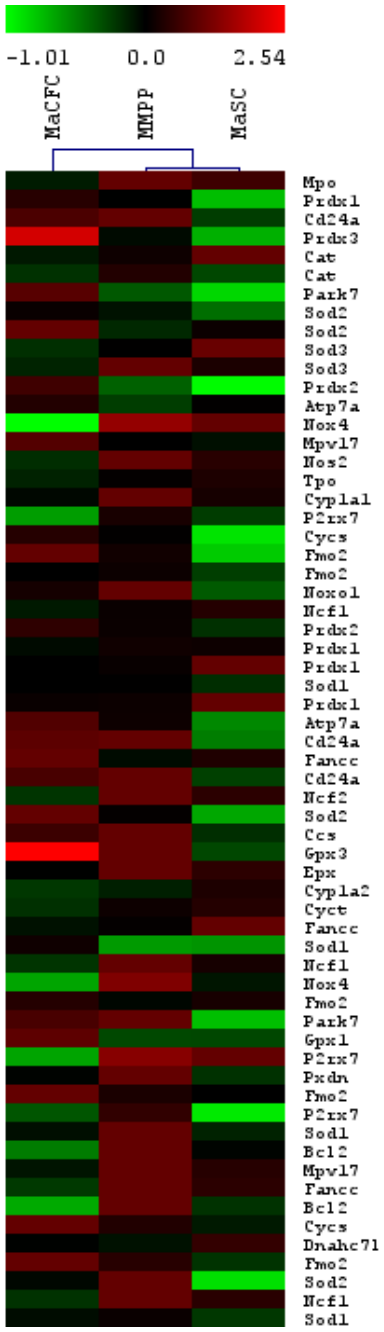



**Figure 4.2: Heat map showing microarray expression data for ROS genes.**

MaSC, MMPP and MaCFC populations were investigated for all known genes that are involved in reducing ROS levels in mouse cells. Clustering analysis of the samples based on centered Pearson correlation with average distance metric. Data for each population is the average expression from 3 replicate arrays.

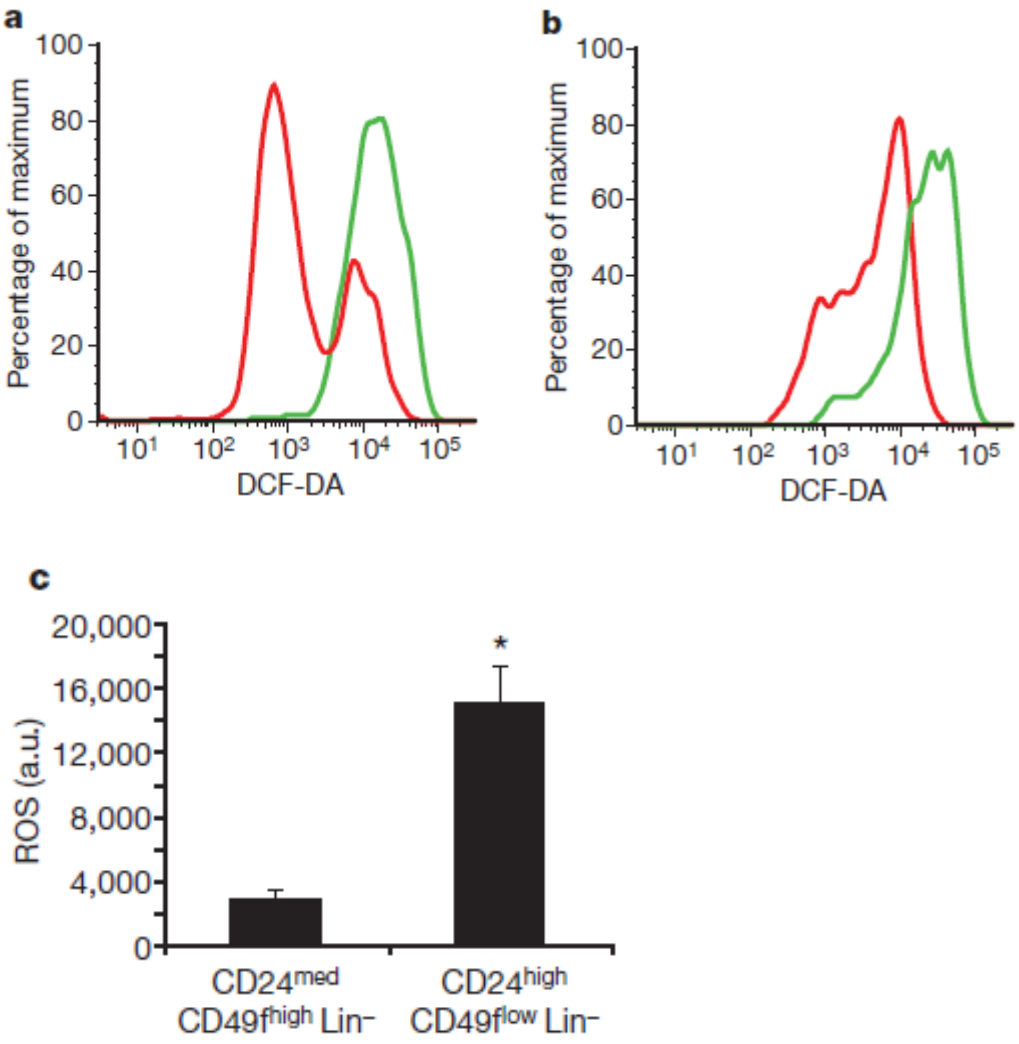



**Figure 4.3: Analysis of ROS levels in normal mammary stem cells and their progeny.**

A, $CD24^{med}$ $CD49f^{hi}Lin^{-}$ mammary cells (mammary stem-cell-containing population; red) and $CD24^{hi}CD49f^{low}Lin^{-}$ mammary cells (progenitor cell-containing population; green) were isolated from C57BL/6J female mice using flow cytometry, and intracellular ROS concentrations were measured by DC-FDA staining. B, As in A, except using 29S1/SvImJ mice. C, The mean and S.E.M. for replicates of A and B are shown (n=6, **p*=0.001). a.u., arbitrary units. These results were representative of the 3 independent experiments performed.

| Phenotype | No. of positive pads / total injections |
|---|---|
| | |
| ROS low | 3/8 |
| | |
| ROS medium | 4/7 |
| | |

**Figure 4.4: Transplantation of MRU-enriched cells based on intracellular levels of ROS.**

$CD24^{med}CD49f^{hi}Lin^{-}$ mammary cells (enriched for mammary repopulating units) were isolated from mammary fat pads from C57Bl/6J female mice and were immediately loaded with 10 μM DCF-DA. Labeled cells were then re-sorted into "ROS-low" and "ROS-mid" sub-populations based on their DCF-DA staining profile (in comparison to that of $CD24^{hi}CD49f^{low}$Lin- progenitor cells, which displayed a "ROS-high" profile). Sorted cells (500-5000 depending on experiment) were injected into cleared mammary fat pads of 21-day-old female C57Bl/6J mice and were scored by wholemount analysis 5–6 weeks later.

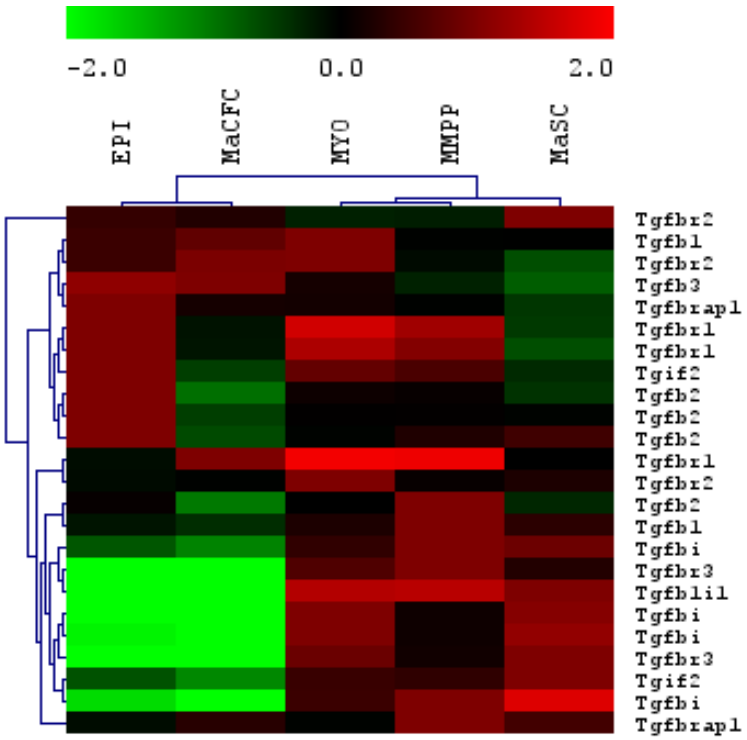
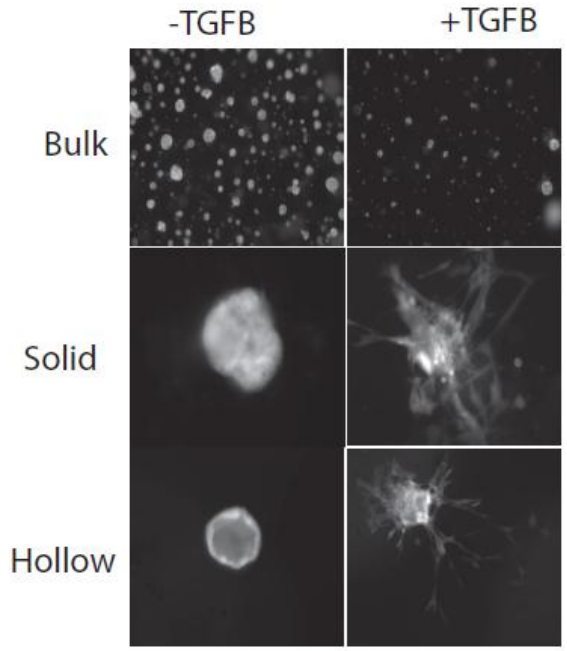



**Figure 4.5: Heat map showing expression of TGFβ ligand and receptors in microarrays.**

Each sample represents the average of 3 replicate arrays. Clustering of samples and genes based on centered Pearson correlation with average distance metric.



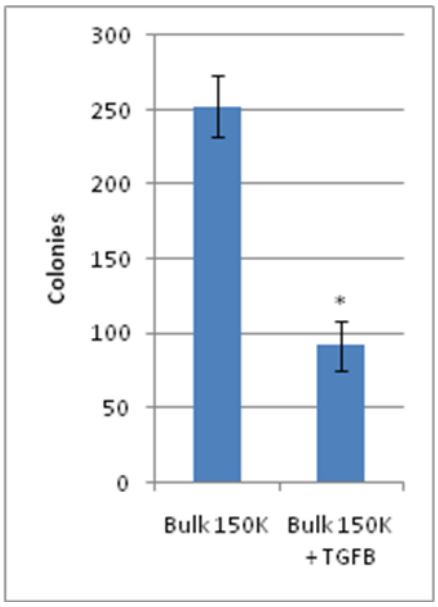




**Figure 4.6: TGFβ1 ligand addition to unsorted mammary cells.**

Top, fluorescence pictures of morphological colony changes observed in unsorted cells when TGFβ1 ligand is added to culture. Bottom, colony forming quantitation of to assess effect of TGFβ1. Data represents three independent experiments, ±S.D. of the mean.

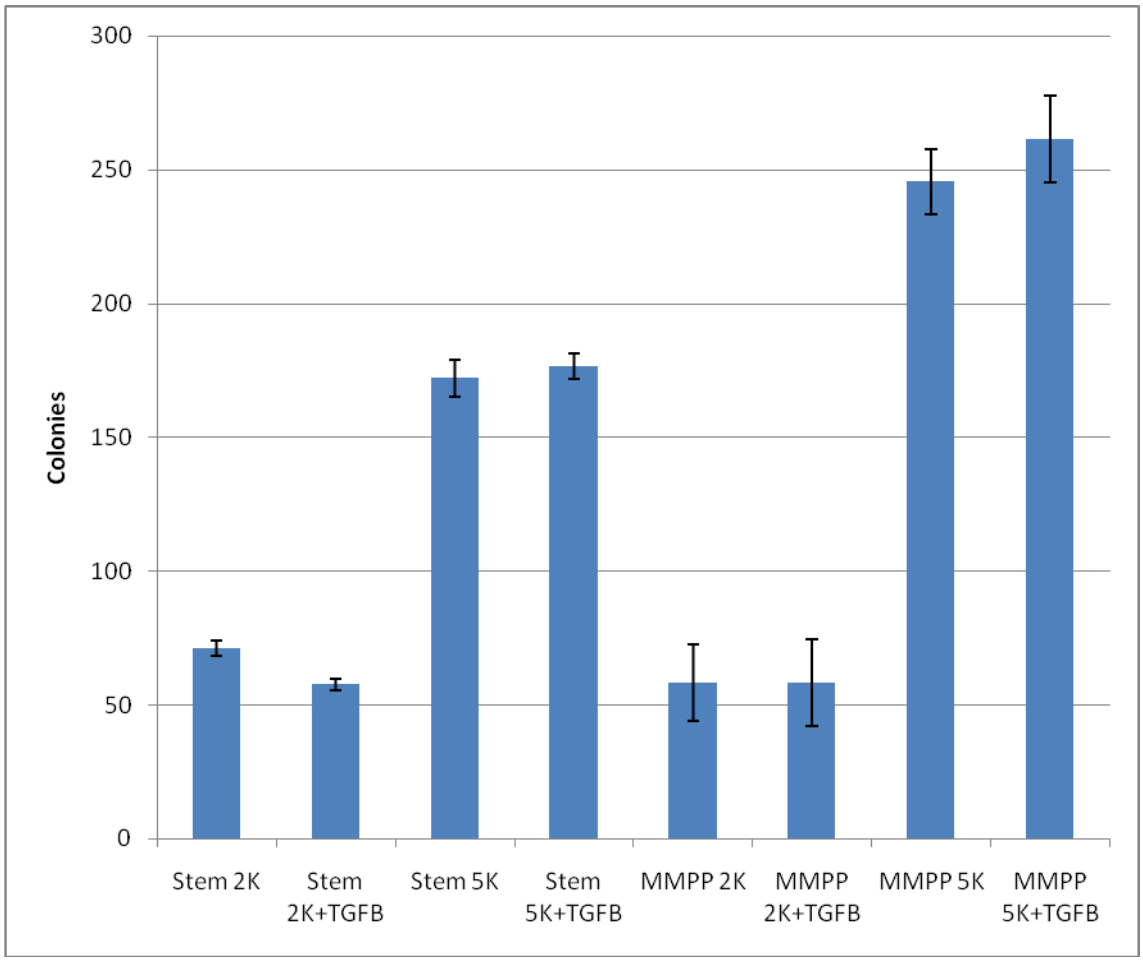



**Figure 4.7: TGFβ1 has no significant effect on colony formation of MaSC or MMPP cells.**

Sorted cells were plated into culture with and without ligand. Colony formation was used as a metric for TGFβ1 effects. Data represents three independent experiments, ±S.D. of the mean.

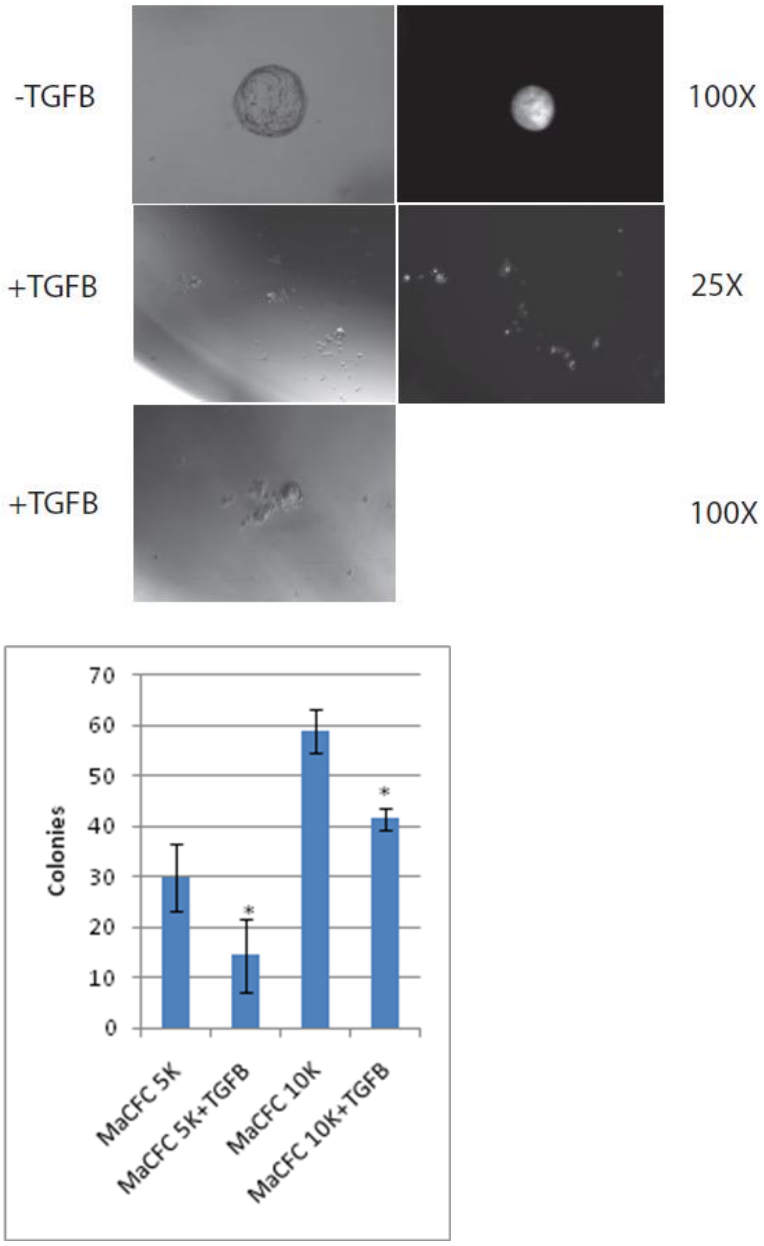



**Figure 4.8: TGFβ1 induces a motility program in MaCFCs.**

Progenitor MaCFC cells' colony formation was assessed when TGFβ1 was added to cultures. Colonies became noticeably smaller due to a lack of proliferation, and a migratory program was initiated in some cells. There was also significant reduction in the number of colonies formed. Data represents three independent experiments, ±S.D. of the mean.

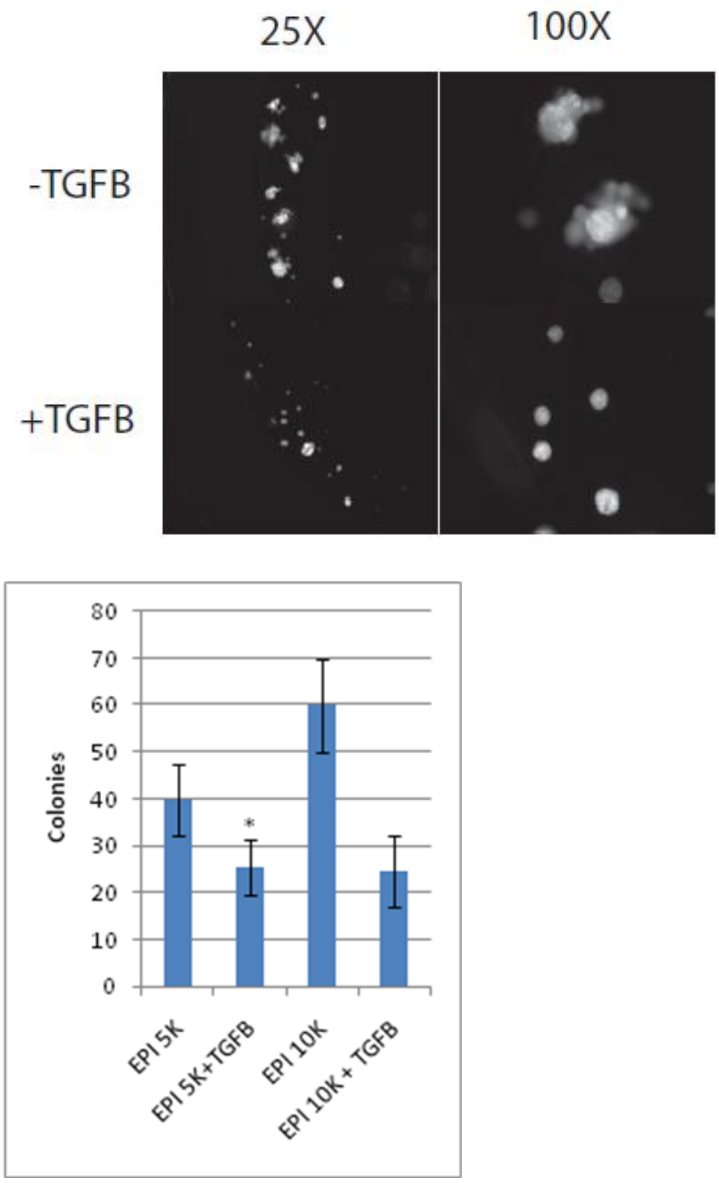



**Figure 4.9: TGFβ1 has a cytostatic effect on EPI cells.**

TGFβ1 addition to cultures of EPI cells resulted in an constrained colony morphology. Also, TGFβ1 significantly reduced colony formation. Data represents three independent experiments, ±S.D. of the mean.

**BIBLIOGRAPHY**


1. Zorov DB, Juhaszova M, Sollott SJ: **Mitochondrial ROS-induced ROS release: an update and review**. *Biochim Biophys Acta* 2006, **1757**:509-517.
2. Griendling KK, FitzGerald GA: **Oxidative stress and cardiovascular injury: Part I: basic mechanisms and in vivo monitoring of ROS**. *Circulation* 2003, **108**:1912-1916.
3. Chakraborti T, Das S, Mondal M, Roychoudhury S, Chakraborti S: **Oxidant, mitochondria and calcium: an overview**. *Cell Signal* 1999, **11**:77-85.
4. Wilhelm J, Vytasek R, Ostadalova I, Vajner L: **Evaluation of different methods detecting intracellular generation of free radicals**. *Mol Cell Biochem* 2009.
5. Panov A, Dikalov S, Shalbuyeva N, Hemendinger R, Greenamyre JT, Rosenfeld J: **Species- and tissue-specific relationships between mitochondrial permeability transition and generation of ROS in brain and liver mitochondria of rats and mice**. *Am J Physiol Cell Physiol* 2007, **292**:C708-718.
6. Atlante A, Calissano P, Bobba A, Azzariti A, Marra E, Passarella S: **Cytochrome c is released from mitochondria in a reactive oxygen species (ROS)-dependent fashion and can operate as a ROS scavenger and as a respiratory substrate in cerebellar neurons undergoing excitotoxic death**. *J Biol Chem* 2000, **275**:37159-37166.
7. Di Giovanni S, Mirabella M, Papacci M, Odoardi F, Silvestri G, Servidei S: **Apoptosis and ROS detoxification enzymes correlate with cytochrome c oxidase deficiency in mitochondrial encephalomyopathies**. *Mol Cell Neurosci* 2001, **17**:696-705.
8. Vertechy M, Cooper MB, Ghirardi O, Ramacci MT: **Antioxidant enzyme activities in heart and skeletal muscle of rats of different ages**. *Exp Gerontol* 1989, **24**:211-218.
9. Jenkins RR, Friedland R, Howald H: **The relationship of oxygen uptake to superoxide dismutase and catalase activity in human skeletal muscle**. *Int J Sports Med* 1984, **5**:11-14.
10. Gonzalez-Lama Z, Feinstein RN: **A soluble peroxidase in heart and skeletal muscle of rat and mouse**. *Proc Soc Exp Biol Med* 1977, **154**:322-324.
11. Harman D: **Aging: a theory based on free radical and radiation chemistry**. *J Gerontol* 1956, **11**:298-300.
12. Diehn M, Cho RW, Lobo NA, Kalisky T, Dorie MJ, Kulp AN, Qian D, Lam JS, Ailles LE, Wong M, et al.: **Association of reactive oxygen species levels and radioresistance in cancer stem cells**. *Nature* 2009.
13. Yao EH, Yu Y, Fukuda N: **Oxidative stress on progenitor and stem cells in cardiovascular diseases**. *Curr Pharm Biotechnol* 2006, **7**:101-108.
14. Terzioglu M, Larsson NG: **Mitochondrial dysfunction in mammalian ageing**. *Novartis Found Symp* 2007, **287**:197-208; discussion 208-113.
15. Singh KK: **Mitochondria damage checkpoint, aging, and cancer**. *Ann N Y Acad Sci* 2006, **1067**:182-190.
16. Huang H, Manton KG: **The role of oxidative damage in mitochondria during aging: a review**. *Front Biosci* 2004, **9**:1100-1117.

17. Naka K, Muraguchi T, Hoshii T, Hirao A: **Regulation of reactive oxygen species and genomic stability in hematopoietic stem cells**. *Antioxid Redox Signal* 2008, **10**:1883-1894.
18. Zhang X, Sejas DP, Qiu Y, Williams DA, Pang Q: **Inflammatory ROS promote and cooperate with the Fanconi anemia mutation for hematopoietic senescence**. *J Cell Sci* 2007, **120**:1572-1583.
19. Tothova Z, Kollipara R, Huntly BJ, Lee BH, Castrillon DH, Cullen DE, McDowell EP, Lazo-Kallanian S, Williams IR, Sears C, et al.: **FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress**. *Cell* 2007, **128**:325-339.
20. Taylor AW: **Review of the activation of TGF-beta in immunity**. *J Leukoc Biol* 2009, **85**:29-33.
21. Haufel T, Dormann S, Hanusch J, Schwieger A, Bauer G: **Three distinct roles for TGF-beta during intercellular induction of apoptosis: a review**. *Anticancer Res* 1999, **19**:105-111.
22. Akhurst RJ, Fitzpatrick DR, Fowlis DJ, Gatherer D, Millan FA, Slager H: **The role of TGF-beta s in mammalian development and neoplasia**. *Mol Reprod Dev* 1992, **32**:127-135.
23. Daniel CW, Robinson S, Silberstein GB: **The role of TGF-beta in patterning and growth of the mammary ductal tree**. *J Mammary Gland Biol Neoplasia* 1996, **1**:331-341.
24. Arteaga CL, Moses HL: **TGF-beta in mammary development and neoplasia**. *J Mammary Gland Biol Neoplasia* 1996, **1**:327-329.
25. Daniel CW, Robinson SD: **Regulation of mammary growth and function by TGF-beta**. *Mol Reprod Dev* 1992, **32**:145-151.
26. Zugmaier G, Lippman ME: **Effects of TGF beta on normal and malignant mammary epithelium**. *Ann N Y Acad Sci* 1990, **593**:272-275.
27. Bierie B, Gorska AE, Stover DG, Moses HL: **TGF-beta promotes cell death and suppresses lactation during the second stage of mammary involution**. *J Cell Physiol* 2009, **219**:57-68.
28. Yang YA, Tang B, Robinson G, Hennighausen L, Brodie SG, Deng CX, Wakefield LM: **Smad3 in the mammary epithelium has a nonredundant role in the induction of apoptosis, but not in the regulation of proliferation or differentiation by transforming growth factor-beta**. *Cell Growth Differ* 2002, **13**:123-130.
29. Ewan KB, Shyamala G, Ravani SA, Tang Y, Akhurst R, Wakefield L, Barcellos-Hoff MH: **Latent transforming growth factor-beta activation in mammary gland: regulation by ovarian hormones affects ductal and alveolar proliferation**. *Am J Pathol* 2002, **160**:2081-2093.
30. Boulanger CA, Wagner KU, Smith GH: **Parity-induced mouse mammary epithelial cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression**. *Oncogene* 2005, **24**:552-560.
31. Kordon EC, McKnight RA, Jhappan C, Hennighausen L, Merlino G, Smith GH: **Ectopic TGF beta 1 expression in the secretory mammary epithelium induces early senescence of the epithelial stem cell population**. *Dev Biol* 1995, **168**:47-61.

32. Jhappan C, Geiser AG, Kordon EC, Bagheri D, Hennighausen L, Roberts AB, Smith GH, Merlino G: **Targeting expression of a transforming growth factor beta 1 transgene to the pregnant mammary gland inhibits alveolar development and lactation**. *Embo J* 1993, **12**:1835-1845.
33. Muraoka-Cook RS, Shin I, Yi JY, Easterly E, Barcellos-Hoff MH, Yingling JM, Zent R, Arteaga CL: **Activated type I TGFbeta receptor kinase enhances the survival of mammary epithelial cells and accelerates tumor progression**. *Oncogene* 2006, **25**:3408-3423.
34. Pierce DF, Jr., Gorska AE, Chytil A, Meise KS, Page DL, Coffey RJ, Jr., Moses HL: **Mammary tumor suppression by transforming growth factor beta 1 transgene expression**. *Proc Natl Acad Sci U S A* 1995, **92**:4254-4258.
35. Bailey JP, Nieport KM, Herbst MP, Srivastava S, Serra RA, Horseman ND: **Prolactin and transforming growth factor-beta signaling exert opposing effects on mammary gland morphogenesis, involution, and the Akt-forkhead pathway**. *Mol Endocrinol* 2004, **18**:1171-1184.
36. Crowley MR, Bowtell D, Serra R: **TGF-beta, c-Cbl, and PDGFR-alpha the in mammary stroma**. *Dev Biol* 2005, **279**:58-72.
37. Gorska AE, Jensen RA, Shyr Y, Aakre ME, Bhowmick NA, Moses HL: **Transgenic mice expressing a dominant-negative mutant type II transforming growth factor-beta receptor exhibit impaired mammary development and enhanced mammary tumor formation**. *Am J Pathol* 2003, **163**:1539-1549.
38. Ito K, Hirao A, Arai F, Matsuoka S, Takubo K, Hamaguchi I, Nomiyama K, Hosokawa K, Sakurada K, Nakagata N, et al.: **Regulation of oxidative stress by ATM is required for self-renewal of haematopoietic stem cells**. *Nature* 2004, **431**:997-1002.
39. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, Wu L, Lindeman GJ, Visvader JE: **Generation of a functional mammary gland from a single stem cell**. *Nature* 2006, **439**:84-88.
40. Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, Li HI, Eaves CJ: **Purification and unique properties of mammary epithelial stem cells**. *Nature* 2006, **439**:993-997.
41. Ghaffari S: **Oxidative stress in the regulation of normal and neoplastic hematopoiesis**. *Antioxid Redox Signal* 2008, **10**:1923-1940.
42. Miyamoto K, Araki KY, Naka K, Arai F, Takubo K, Yamazaki S, Matsuoka S, Miyamoto T, Ito K, Ohmura M, et al.: **Foxo3a is essential for maintenance of the hematopoietic stem cell pool**. *Cell Stem Cell* 2007, **1**:101-112.
43. Jang YY, Sharkis SJ: **A low level of reactive oxygen species selects for primitive hematopoietic stem cells that may reside in the low-oxygenic niche**. *Blood* 2007, **110**:3056-3063.
44. Hosokawa K, Arai F, Yoshihara H, Nakamura Y, Gomei Y, Iwasaki H, Miyamoto K, Shima H, Ito K, Suda T: **Function of oxidative stress in the regulation of hematopoietic stem cell-niche interaction**. *Biochem Biophys Res Commun* 2007, **363**:578-583.
45. Junn E, Lee KN, Ju HR, Han SH, Im JY, Kang HS, Lee TH, Bae YS, Ha KS, Lee ZW, et al.: **Requirement of hydrogen peroxide generation in TGF-beta 1 signal transduction in human lung fibroblast cells: involvement of hydrogen**

**peroxide and Ca2+ in TGF-beta 1-induced IL-6 expression**. *J Immunol* 2000, **165**:2190-2197.
46. Gooch JL, Gorin Y, Zhang BX, Abboud HE: **Involvement of calcineurin in transforming growth factor-beta-mediated regulation of extracellular matrix accumulation**. *J Biol Chem* 2004, **279**:15561-15570.
47. Okabe M, Ikawa M, Kominami K, Nakanishi T, Nishimune Y: **'Green mice' as a source of ubiquitous green cells**. *FEBS Lett* 1997, **407**:313-319.
48. Stingl J: **Detection and analysis of mammary gland stem cells**. *J Pathol* 2009, **217**:229-241.
49. Ito K, Takubo K, Arai F, Satoh H, Matsuoka S, Ohmura M, Naka K, Azuma M, Miyamoto K, Hosokawa K, et al.: **Regulation of reactive oxygen species by Atm is essential for proper response to DNA double-strand breaks in lymphocytes**. *J Immunol* 2007, **178**:103-110.
50. Wakaumi M, Ishibashi K, Ando H, Kasanuki H, Tsuruoka S: **Acute digoxin loading reduces ABCA8A mRNA expression in the mouse liver**. *Clin Exp Pharmacol Physiol* 2005, **32**:1034-1041.
51. Bertucci F, Finetti P, Cervera N, Charafe-Jauffret E, Buttarelli M, Jacquemier J, Chaffanet M, Maraninchi D, Viens P, Birnbaum D: **How different are luminal A and basal breast cancers?** *Int J Cancer* 2009, **124**:1338-1348.
52. Cho RW, Wang X, Diehn M, Shedden K, Chen GY, Sherlock G, Gurney A, Lewicki J, Clarke MF: **Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors**. *Stem Cells* 2008, **26**:364-371.
53. Shin SR, Sanchez-Velar N, Sherr DH, Sonenshein GE: **7,12-dimethylbenz(a)anthracene treatment of a c-rel mouse mammary tumor cell line induces epithelial to mesenchymal transition via activation of nuclear factor-kappaB**. *Cancer Res* 2006, **66**:2570-2575.

# CHAPTER 5

## CONCLUSIONS AND FUTURE DIRECTIONS

In my thesis work, I began with a classic stem cell marker study on the murine mammary system. During my initial studies, the enrichment of the mammary stem cell based on the surface marker phenotype $CD24^{med}CD49f^{hi}CD29^{hi}Sca\text{-}1^{lo}$ was published[1,2]. This work definitively proved the existence of the mammary stem cell using FACS in combination with the power of *in vivo* transplantation into mammary fat pads cleared of endogenous ductal epithelium. These results were encouraging for my work, as it provided me a platform to further advance current isolation strategies. During this same period, our lab successfully isolated a tumor initiating population for the *MMTV-Wnt1* mouse model of breast cancer based on the $Thy\text{-}1^{+}CD24^{+}$ phenotype[3]. Viewing tumors as aberrant version of normal tissue, we reasoned that Thy-1 and CD24 expression may prove useful in isolating progenitor cells for normal breast. As CD24 was previously shown to enrich for mammary stem cells[4], I began a rigorous study of Thy-1 expression in the normal mouse mammary gland. Previous work on Thy-1 revealed the protein was expressed on thymic epithelium[5], but more importantly was found on mouse hematopoietic stem cells[6] and basal/myoepithelial cells in the rat mammary gland[7]. Interestingly, Thy-1 knockout mice were fertile and seemingly normal[8].

Monitored behavior observation indicated these mice failed to learn social cues from their littermates due to defects in GABA receptor signaling. Also, monoclonal Thy-1.1 blocking antibody causes mesangioproliferative glomerulonephritis in rats[9], showing activation of Thy-1 mediated signaling has additional brain effects. These and other subsequent studies collectively showed Thy-1 had a role in T-cell activation, apoptosis, tumor suppression, fibrosis and apoptosis[10,11]. Even though Thy-1 was well studied in a number of organ systems, little was known about its role, if any, in the mouse mammary epithelium.

As described in chapter 1, my initial work assessed how Thy-1 was expressed in the mammary gland. Assessing protein or mRNA levels by classical methods such as western and northern blot would not provide specific information about individual cell types, so I used flow cytometry instead. As mentioned earlier, many groups had shown flow cytometry could relate information of protein presence on specific cell type when used in combinations with other known markers. I found Thy-1 was differentially expressed on dissociated mammary cells; importantly, Thy-1 was expressed on a fraction of the $CD24^{med}CD49f^{hi}$ population that had been shown to be enriched for mammary stem cells. Also, I found Sca-1 was expressed on a small percentage of $CD24^{med}CD49f^{hi}$ cells at a high level, confirming previous observations. Neither $Thy\text{-}1^{+}CD24^{med}CD49f^{hi}$ cells nor $Thy\text{-}1^{-}CD24^{med}CD49f^{hi}$ cells were adept at making colonies when plated on NIH3T3 feeder cells, a classical mammary differentiation assay. However, myoepithelial cells were known to be poor colony forming cells, thus I reasoned the mammary stem cell had a myoepithelial phenotype. Immunohistochemistry showed Thy-1 localized to the basal cells in ducts, in the same region as myoepithelial cells. I then proceeded to test the

potential of Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ and Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells to make ductal epithelium in vivo. I found duct forming ability was significantly enriched in the Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ population, suggesting Thy-1 was present on mammary stem cells. I did observe some duct forming ability in the Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells. A defining trait of all stem cells is self-renewal, or the process of one stem cell making another one. To measure the ability of Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ epithelium to self-renew, I removed pieces of primary transplants and performed secondary transplants. Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ primary epithelium had an increased ability to self-renew compared to Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ epithelium, confirming Thy-1 did enrich for stem cells. In addition, Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ cells were able to reproduce all of the heterogeneity of mammary ducts, including Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells. When comparing my results to the prototypical stem cell system in the blood, I found the Thy-1$^{-}$CD24$^{med}$CD49f$^{hi}$ cells were analogous to the multipotent progenitor cells that have no self-renewal ability but can transiently reconstitute the blood system. Therefore, I proposed the Thy-1$^{+}$CD24$^{med}$CD49f$^{hi}$ cells were the mammary stem cells (MaSC) and the Thy-1$^{-}$ CD24$^{med}$CD49f$^{hi}$ cells were the mammary multipotent progenitor cells (MMPP).

The recent elucidation of the cellular hierarchy in the mammary system has value and application to a wide variety of important topics. Of particular interest to me is the question of which cells are the most likely to turn into cancer[12]. When viewing cancers in a developmental perspective, it has been proposed that stem and progenitors cells may be the most common targets of transformation since they are the most long lived populations in any somatic tissue and go through the most cell division. These two characteristics make them more likely to acquire mutations compared to more

differentiated , shorter lived cells. Also, many tumors have considerable heterogeneity and stages of differentiation represented in the whole population of cancerous cells, and more aggressive breast tumors usually are poorly differentiated and have a basal phenotype. Therefore, I believe the MaSC and the MMPP cells would produce tumors at greater frequency and that are more aggressive based on the introduction of oncogene(s). This hypothesis may also explain why tumorigenic cells from *MMTV-Wnt1* tumors have a more basal phenotype compared to non-tumorigenic tumor cells[3].

In chapter 2, I analyzed the global gene expression of isolated mammary cell populations by microarrays. The populations I interrogated were the MaSC, MMPP, MaCFC, MYO and EPI cells from virgin C57BL/6 mice. These arrays profiled the most highest numbers of progenitor cells recorded[13], giving us better signal and resolution. My goals were to understand how these progenitor populations were related and to identify genes that were only expressed by the MaSCs. Unsupervised hierarchical clustering showed the MaCFC population had similar phenotype to differentiated luminal cells and the MaSC and MMPP populations were more similar to the myoepithelial phenotype. When unsupervised clustering was applied to analyze all cytokeratins' expression, the results were the same as general clustering, except this method gave information about two relatively unstudied keratins, Krt23 and Krt80. Keratin 23 was expressed in the MMPP and MYO arrays and Krt80 was expressed by the MaSC and EPI arrays. The keratin analysis also confirmed previous descriptive studies that showed MaSCs and MMPPs expressed the myoepithelial keratins 5 and 14 while the MaCFCs expressed keratin 8 and 18. Interestingly, the MaCFCs also expressed keratin 6, a molecule that had been proposed as a progenitor marker[14-16]. In addition, MaCFCs

and EPI arrays had expression of keratin 19, a poorly understood keratin that was thought to mark a subset of luminal cells[17,18]. Real-time PCR validation confirmed my array results, although I wasn't able to detect keratin 6 expression in any isolated population. Regardless, the real-time data also showed MaSCs expressed myoepithelial keratins as well as a low amount of keratin 18, suggesting the stem cells retain some markers of both differentiated lineages. I also found some expression differences between Thy-$1^{+}$MaCFCs and Thy-$1^{-}$MaCFCs, suggesting there Thy-1 may also prove useful in segregating these early progenitor cells into multiple intermediate progenitors. Encouraged by these results, I measured $p21^{cip1/waf1}$ and Tbx3 expression by real-time PCR and found $p21^{cip1/waf1}$ was preferentially expressed in the MaSCs whereas Tbx3 was not. These data further confirmed my microarray data as well as providing new insight into the defects involved in Ulnar Mammary Syndrome. I used a novel *ex vivo* analysis system to check protein levels of some of the better defined keratins in the different mammary populations. These results were similar to my real-time PCR expression data, suggesting post-transcriptional modification was not an important regulatory system that changed how I interpreted the keratin data from the microarrays.

Through my microarray analysis, I took the opportunity to explore keratin expression in histology and cultured cells. After staining sections to localize where KRT8, KRT14, KRT19 and KRT6 express in the mature ductal epithelium. I found, as previous groups had described[19,20], KRT8 is expressed by luminal cells and KRT14 was expressed by myoepithelial cells. I also showed KRT19 and KRT6 was expressed by a subset of luminal cells. These data may shed light in the location of the MaCFC progenitors, as they were the only population that expressed Krt6 in microarray analysis.

Furthermore, my histology was the first evidence of localized Krt19 expression in the mouse mammary gland, although it had been previously postulated. An interesting future direction would be to look at how pregnancy affects the Krt19 populations, assuming that is does mark an alveolar progenitor population. Specifically, linking Whey Acidic Proein (*Wap*) expression with keratin 19 at the single cell level could provide proof that Krt19 is a marker that could be used to distinguish mammary stem cells from alveolar progenitors. Also, monitoring Krt6 expression in combination with Krt23 and Krt80 may help to localize the stem cells.

Continuing my microarray analysis, I was able to distinguish a set of genes that were specifically expressed in the MaSC. This was the first report of such a gene signature. I did not explore in detail which genes were specifically downregulated in the MaSCs, as this work is ongoing. I have found that the expression of genes is easier to investigate (via loss-of-function studies) rather than the absence of expression. The genes that were specific to the MaSCs provided intriguing suggestions about how mammary stem cells behave in their natural state as well as their predisposition for activating transcriptional programs based on ion channel signaling information. Further work would involve identifying the specific ion channels that are important for this physiologic activity, and then manipulating this signaling to elicit specific stem cell responses. Such studies could provide important information about how the mammary tissue is maintained throughout development into adulthood, and has implications for oncogenesis (i.e. if ion channels can be controlled to prevent aberrant proliferation of stem cells).

In chapter 3, I investigated two distinct mechanisms that had been speculated were involved in stem cell fate decisions. The first study was based on the observation

that hematopoietic stem cells have lower reactive oxygen species (ROS) that further differentiated progeny[21]. This observation was also true of the mammary stem cell population. I also showed the populations enriched for mammary stem cells could be divided into ROS high and low containing cells. The cells that had the lowest amount of intracellular ROS had an increased capacity to produce ductal epithelium in transplantation, suggesting stem cells may share the ability to reduce their ROS levels as a method of protecting themselves from genotoxic substrates. My studies provide a good platform to evaluate which specific genes are involved in this mechanism, since cells can use enzymatic (i.e. Catalase) or non-enzymatic (efflux pumps) methods of reducing their ROS content. In addition, since cancer stem cells also share this property this may be evidence of their origin in normal tissue.

Finally, I took advantage of the power of the population isolation strategies from my previous work to demonstrate how TGFβ signaling affects mammary cells. I found different mammary populations had specific reactions to exogenous ligand, with early progenitors inducing a migratory program. Surprisingly, there was little effect on MaSCs and MMPPs, which had been previously speculated[22]. My work does not rule out these possibilities, since MaSCs and MMPPs may have autocrine TGFβ signaling loops. Therefore, loss-of-function studies with TGFβ receptors could prove useful in determining the specificity of the system. In addition, there are many members of the TGF superfamily, and it is reasonable to expect a combination of ligand-mediated signals would be needed to affect the stem cells. However, the complexity of these studies may hinder the accurate interpretations of any positive results. Nevertheless, given the

importance of TGFβ signaling in a wide range of cellular behaviors, further work using these isolation strategies may be key to deciphering TGFβ effects.

In conclusion, my work has provided many interesting and novel insights into the murine mammary epithelial cellular hierarchy. This thesis comprises both functional and descriptive work that profiled the mammary system in ways not achieved previously. I believe the rich information gleaned from my microarray data may be used to generate hypotheses about how mammary stem cells behave, and perhaps more importantly, which systems may be corrupted during oncogenesis.

## BIBLIOGRAPHY


1. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, Wu L, Lindeman GJ, Visvader JE: **Generation of a functional mammary gland from a single stem cell**. *Nature* 2006, **439**:84-88.
2. Stingl J: **Detection and analysis of mammary gland stem cells**. *J Pathol* 2009, **217**:229-241.
3. Cho RW, Wang X, Diehn M, Shedden K, Chen GY, Sherlock G, Gurney A, Lewicki J, Clarke MF: **Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors**. *Stem Cells* 2008, **26**:364-371.
4. Sleeman KE, Kendrick H, Ashworth A, Isacke CM, Smalley MJ: **CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells**. *Breast Cancer Res* 2006, **8**:R7.
5. Raedler A, Raedler E, Scholz KU, Arndt R, Thiele HG: **The intrathymic microenvironment: expression of lectin receptors and lectin-like molecules of differentiation antigens and MHC gene products**. *Thymus* 1983, **5**:311-326.
6. Morrison SJ, Hemmati HD, Wandycz AM, Weissman IL: **The purification and characterization of fetal liver hematopoietic stem cells**. *Proc Natl Acad Sci U S A* 1995, **92**:10302-10306.
7. Rudland PS, Warburton MJ, Monaghan P, Ritter MA: **Thy-1 antigen on normal and neoplastic rat mammary tissues: changes in location and amount of antigen during differentiation of cultured stem cells**. *J Natl Cancer Inst* 1982, **68**:799-811.
8. Simon PD, McConnell J, Zurakowski D, Vorwerk CK, Naskar R, Grosskreutz CL, Dreyer EB: **Thy-1 is critical for normal retinal development**. *Brain Res Dev Brain Res* 1999, **117**:219-223.
9. Schlondorff D: **Roles of the mesangium in glomerular function**. *Kidney Int* 1996, **49**:1583-1585.
10. Barker TH, Hagood JS: **Getting a grip on Thy-1 signaling**. *Biochim Biophys Acta* 2008.
11. Rege TA, Hagood JS: **Thy-1, a versatile modulator of signaling affecting cellular adhesion, proliferation, survival, and cytokine/growth factor responses**. *Biochim Biophys Acta* 2006, **1763**:991-999.
12. Lobo NA, Shimono Y, Qian D, Clarke MF: **The biology of cancer stem cells**. *Annu Rev Cell Dev Biol* 2007, **23**:675-699.
13. Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, Li HI, Eaves CJ: **Purification and unique properties of mammary epithelial stem cells**. *Nature* 2006, **439**:993-997.
14. Luo M, Fan H, Nagy T, Wei H, Wang C, Liu S, Wicha MS, Guan JL: **Mammary epithelial-specific ablation of the focal adhesion kinase suppresses mammary tumorigenesis by affecting mammary cancer stem/progenitor cells**. *Cancer Res* 2009, **69**:466-474.
15. Stingl J, Raouf A, Emerman JT, Eaves CJ: **Epithelial progenitors in the normal human mammary gland**. *J Mammary Gland Biol Neoplasia* 2005, **10**:49-59.

16. Welm B, Behbod F, Goodell MA, Rosen JM: **Isolation and characterization of functional mammary gland stem cells**. *Cell Prolif* 2003, **36 Suppl 1**:17-32.
17. Bartek J, Bartkova J, Taylor-Papadimitriou J: **Keratin 19 expression in the adult and developing human mammary gland**. *Histochem J* 1990, **22**:537-544.
18. Clarke RB, Spence K, Anderson E, Howell A, Okano H, Potten CS: **A putative human breast stem cell population is enriched for steroid receptor-positive cells**. *Dev Biol* 2005, **277**:443-456.
19. Alvi AJ, Clayton H, Joshi C, Enver T, Ashworth A, Vivanco MM, Dale TC, Smalley MJ: **Functional and molecular characterisation of mammary side population cells**. *Breast Cancer Res* 2003, **5**:R1-8.
20. Vos JH, van den Ingh TS, Misdorp W, Molenbeek RF, van Mil FN, Rutteman GR, Ivanyi D, Ramaekers FC: **Immunohistochemistry with keratin, vimentin, desmin, and alpha-smooth muscle actin monoclonal antibodies in canine mammary gland: malignant mammary tumours**. *Vet Q* 1993, **15**:96-102.
21. Ito K, Hirao A, Arai F, Takubo K, Matsuoka S, Miyamoto K, Ohmura M, Naka K, Hosokawa K, Ikeda Y, et al.: **Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells**. *Nat Med* 2006, **12**:446-451.
22. Li P, Barraclough R, Fernig DG, Smith JA, Rudland PS: **Stem cells in breast epithelia**. *Int J Exp Pathol* 1998, **79**:193-206.