Oncogene (2010) 29, 2753–2759

www.nature.com/onc

REVIEW

Transcription factor regulatory networks in mammary epithelial development and tumorigenesis

PM Siegel and WJ Muller

Departments of Biochemistry and Medicine, Goodman Cancer Centre, McGill University, Montreal, Quebec, Canada

The mouse mammary gland is composed of three epithelial cell types, which include ductal, alveolar and myoepithelial cells. A hierarchy in which the mammary stem cell compartment gives rise to progressively restricted progenitors that ultimately form the luminal (ductal and alveolar) and myoepithelial lineages is now emerging. Although very little is known about the mechanisms controlling the differentiation of the myoepithelial cell lineage, a growing body of work reveals that the luminal cell fate is specified by a network of transcription factors. The precise roles of specific transcription factors in promoting differentiation of luminal progenitors into ductal or alveolar cells are now being elucidated. This review will discuss the importance of these recent observations and place them within the context of other transcription factor networks involved in mammary gland development and tumorigenesis.
Oncogene (2010) **29,** 2753–2759; doi:10.1038/onc.2010.43; published online 29 March 2010

Keywords: mammary gland development; transcription factor networks; STAT; Elf; GATA-3; breast cancer

Introduction

The isolation of epithelial stem cells from both mouse (Shackleton *et al*., 2006; Stingl *et al*., 2006a) and human (Ginestier *et al*., 2007; Eirew *et al*., 2008; Lim *et al*., 2009) mammary glands, and the assessment of their ability to regenerate this tissue *in vivo*, has begun to build the framework for a differentiation hierarchy in the adult mammary gland (Stingl *et al*., 2006b; Molyneux *et al*., 2007; Asselin-Labat *et al*., 2008; Visvader, 2009). In the mouse mammary gland, a single stem cell that is characterized as $CD29^{hi}/CD49f^{hi}/CD24^{+/mod}/Sca\text{-}1^{-}$ can give rise to all epithelial cell types required to produce a functional mammary gland and possesses self-renewal properties (Shackleton *et al*., 2006; Stingl *et al*., 2006a). It is believed that this mammary stem cell gives rise to a common progenitor that then splits into two lineages, which includes the myoepithelial progenitors that produce myoepithelial cells and the luminal progenitors that generate both ductal luminal epithelial cells and alveolar luminal epithelial cells (Asselin-Labat *et al*., 2008).

As in other cell compartments, such as the hematopoietic system, the differentiation potential of mammary epithelial cells within this hierarchical structure appears to be controlled by the action of specific transcription factors. Previous work has demonstrated an important role for steroid hormone transcription factors in regulating distinct phases of mammary gland development, with the estrogen receptor having critical roles in ductal elongation and the progesterone receptor for directing proper lobuloalveolar development during pregnancy (Brisken, 2002; Sternlicht *et al*., 2006). In addition to this class of transcription factors, the CCAAT/enhancer-binding protein, C/EBPbeta, also has a critical function in promoting proper lobuloalveolar development in the mammary gland (Grimm and Rosen, 2003; Lamarca *et al*., 2010). Although these factors are clearly necessary for the proper development of the virgin mammary gland ductal tree and subsequent development during pregnancy, it is not yet clear how these proteins control the specification of distinct cell types within the emerging differentiation hierarchy of the mammary gland. In this respect, recent work has begun to focus on elucidating the roles and interactions of several key transcription factors during normal mammary epithelial differentiation of specific mammary epithelial populations (Kouros-Mehr *et al*., 2006; Asselin-Labat *et al*., 2007; Oakes *et al*., 2008; Choi *et al*., 2009; Yamaji *et al*., 2009). In addition to terminal differentiation, another set of transcription factors are involved in the involution phase of mammary gland development, notably STAT3. This review will focus on the roles played by several of these; including STAT5, STAT3, Elf-5 and GATA-3 in both mammary gland development and control of a differentiation hierarchy as well as their potential roles in breast cancer initiation and progression.

STAT5A has a critical role in the establishment of luminal alveolar progenitor cells

Previous work from Hennighausen group had demonstrated that targeted disruption of STAT5A resulted in

Correspondence: Dr WJ Muller, Departments of Biochemistry and Medicine, Goodman Cancer Centre, McGill University, 1160 Avenue des Pins Ouest, Room 516, Montreal, Quebec, Canada H3A 1A3.
E-mail: william.muller@mcgill.ca
Received 30 October 2009; revised 11 January 2010; accepted 19 January 2010; published online 29 March 2010

impaired alveolar differentiation during pregnancy and a failure to lactate (Liu *et al*., 1997; Miyoshi *et al*., 2001). To better understand the nature of this developmental block, a conditional STAT5A/5B allele was generated to permit mammary gland-specific ablation of both these key transcription factors following mammary epithelial expression of Cre recombinase (Yamaji *et al*., 2009). Mammary epithelial disruption of both STAT5A and STAT5B did not negatively impact the mammary stem cell population or their ability to contribute to normal ductal development. Loss of STAT5A/5B did result in a reduced number of CD61$^+$ luminal progenitor cells, but this population did not increase during pregnancy. This led the authors to conclude that two types of CD61$^+$ luminal progenitors exist, those destined to form ductal epithelium and those that undergo alveolar differentiation. The loss of STAT5A/5B did not affect the population of CD61$^+$ luminal ductal progenitors, but caused a reduction in CD61$^+$ luminal alveolar progenitor cells. Thus, transplants of STAT5A/5B-deficient stem cell-enriched populations could reconstitute a normal ductal tree but failed to generate functional alveoli during pregnancy. Strikingly, normal alveolar development could be restored upon inducible expression of STAT5A. Collectively these data argue that STAT5A is both necessary and sufficient for specification of a luminal alveolar progenitor population. These results now add additional information regarding how specific transcription factors control the differentiation potential of mammary epithelial cells during mammary gland development and builds on previous work in this field (Figure 1).

GATA-3 specifies luminal cell fate in the mammary gland

Another class of transcription factors that has recently been implicated in luminal development is GATA-3. GATA-3 belongs to a family of zinc-finger transcription factors that have important roles in cell-fate specification and has recently emerged as a critical determinant of luminal epithelial cell differentiation within the mammary gland (Naylor and Ormandy, 2007; Kouros-Mehr *et al*., 2008b). GATA-3 expression is localized to the initial sites wherein mammary buds form, implying a role for this factor in the

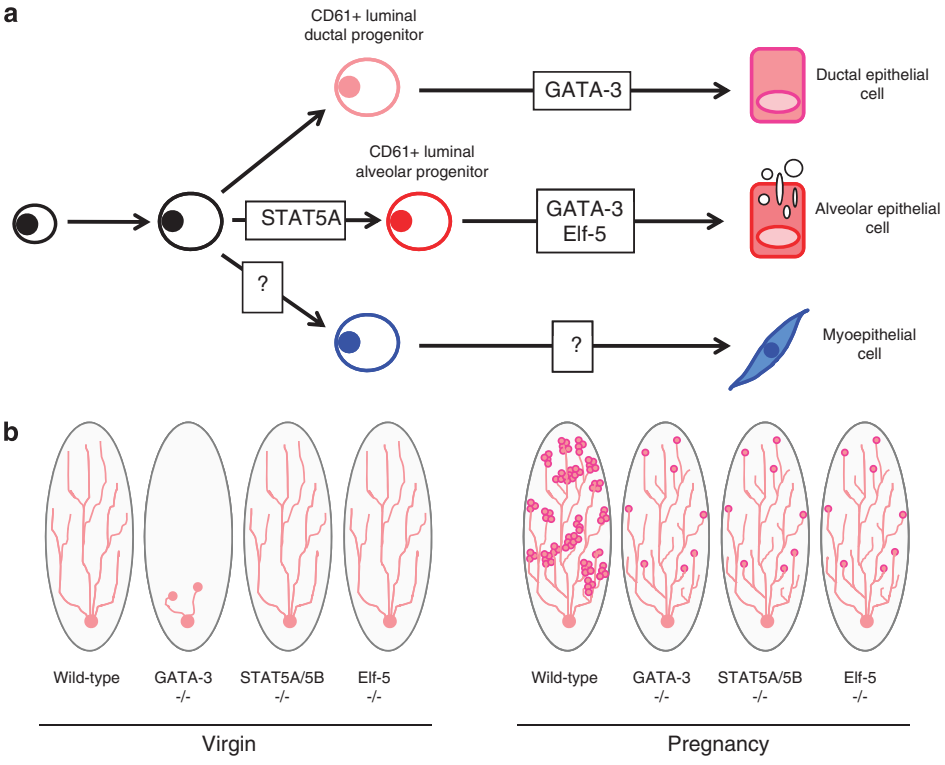

Figure 1 Transcription regulatory network in mammary gland development. (**a**) Schematic illustrating the roles of specific transcription factors in mammary gland development. GATA-3 is involved in generation of both ductal mammary epithelial cells and fully differentiated alveolar epithelial cells. STAT5A is involved in the specification of CD61$^+$ alveolar progenitors but is dispensable for the genesis of CD61$^+$ luminal ductal progenitor cells. Like GATA-3, Elf-5 is required for the induction of the terminally differentiated alveolar epithelial cell population. The question marks denote the lack of information regarding the transcription factor(s) that promote both the specification and differentiation of myoepithelial cells. (**b**) Schematic illustrating the mammary gland phenotypes of virgin and lactating glands derived from ablating the indicated transcription factors. Deletion of GATA-3 early in mammary gland development (MMTV–Cre) results in a severe defect in mammary gland development in virgin animals, whereas deletion of STAT5A/5B and Elf-5 has little impact on ductal outgrowth and branching morphogenesis. In contrast, deletion of GATA-3 (WAP–Cre), STAT5A/5B and Elf-5 have a major impact on the development of terminally differentiated alveolar cells during pregnancy.

mammary primordium (Asselin-Labat *et al*., 2007). Deletion of GATA-3 at this early stage, through the expression of K14-Cre, resulted in the failure to develop mammary placodes and a subsequent lack of a mammary tree in the majority of cases (Asselin-Labat *et al*., 2007).

At later stages of mammary development, GATA-3 expression is restricted to luminal epithelial cells and is absent from the myoepithelial compartment of the developing and adult mammary gland (Kouros-Mehr *et al*., 2006; Asselin-Labat *et al*., 2007). Conditional loss of GATA-3 in the luminal mammary epithelial compartment, by MMTV–Cre-mediated excision of a conditional GATA-3 allele, resulted in the impairment of ductal elongation (Kouros-Mehr *et al*., 2006; Asselin-Labat *et al*., 2007). This result demonstrates a role for GATA-3 in establishing progenitor cells that form the luminal ductal epithelium. If GATA-3 was deleted at a later developmental stage, through the use of WAP–Cre mice, lobuloalveolar development was impaired during pregnancy, resulting in a lactation-deficient phenotype. Analysis of the mammary epithelial compartment in these animals revealed an increased number of luminal cells that lacked differentiation markers, such as caseins (Kouros-Mehr *et al*., 2006; Asselin-Labat *et al*., 2007). Excision of GATA-3, either by MMTV–Cre or WAP–Cre, led to an expansion of the $CD61^+$ progenitor cell population, indicating a role of GATA-3 in the specification of luminal epithelial cells. Conversely, GATA-3 overexpression in mammary stem cell-enriched populations was sufficient to induce the expression of luminal differentiation markers, including WAP and β-casein (Asselin-Labat *et al*., 2007). These data position GATA-3 as a key regulatory factor in the differentiation of mammary progenitor cells; both in the formation of mature ductal epithelial cells during arborization of the mammary fat pad and in subsequent alveolar differentiation during pregnancy.

Elf-5 controls alveolar differentiation

The ETS family transcription factor, Elf-5, has also recently been shown to have a pivotal role in lobuloalveolar differentiation during pregnancy (Oakes *et al*., 2008). In contrast to GATA-3, loss of Elf-5 has no impact on ductal elongation and branching in mammary epithelial explantation experiments. However, lobuloalveolar development was severely impaired in Elf-$5^{-/-}$ transplants. Conversely, forced expression of Elf-5 in luminal epithelial cells of virgin mice resulted in decreased proliferation, impaired ductal outgrowth and precocious differentiation/secretion within the terminal end buds. In Elf-$5^{-/-}$ transplants, an increase in $CD61^+$ luminal progenitor cells was observed, which positions Elf-5 as an important factor within the luminal alveolar lineage (Oakes *et al*., 2008). More recently, the importance of Elf-5 in lobuloalveolar differentiation was supported through the analysis of Elf-5 conditional knockout mice (Choi *et al*., 2009). Like STAT5A-deficient mammary glands, mammary-specific disruption of Elf-5 led to a complete block in alveolar differentiation (Choi *et al*., 2009). Taken together, these studies support Elf-5 as a master regulator of the alveolar switch that occurs during pregnancy.

Building a transcription factor network in mammary gland development

The analysis of mammary gland phenotypes derived from the loss of specific transcription factors has begun to shed light on how lineage commitment during mammary gland differentiation is controlled. A clear relationship exists between STAT5A and Elf-5 during the commitment and differentiation of luminal alveolar cells. Recent work demonstrates that deletion of STAT5A/5B in the mammary glands results in a severe defect in lobuloalveolar differentiation without effects on ductal outgrowth (Yamaji *et al*., 2009). Similarly, Elf-$5^{-/-}$ glands are also capable of normal ductal elongation and branching morphogenesis but fail to undergo proper lobuloalveolar differentiation (Oakes *et al*., 2008; Choi *et al*., 2009). On its face, these results suggest that STAT5A/5B and Elf-5 control similar aspects of lobuloalveolar differentiation. Indeed, the presence of STAT transcription factor-binding sites have been identified within the Elf-5 proximal promoter and Elf5 expression was lost in STAT5A/5B-deficient glands (Yamaji *et al*., 2009), arguing that STAT5A lies immediately upstream of Elf-5 in a transcription factor hierarchy. This regulation is likely more complex, given that the block in alveolar differentiation in Elf-5 null mammary glands may, in part, be attributed to decreased transcriptional activation of STAT5 through Elf-5 transcription binding sites within the STAT5A proximal promoter (Choi *et al*., 2009). However, one important difference between the phenotypes observed in STAT5A/5B and Elf-5 deficient mammary glands is that the Elf-5 deficient glands exhibit an increase in the $CD61^+$ luminal progenitor population (Oakes *et al*., 2008); whereas the STAT5A/5B-deficient glands exhibit a dramatic deficiency in $CD61^+$ progenitor cells. One possible interpretation of this phenotype is that STAT5A is required for the induction of $CD61^+$ luminal progenitors that are destined to form alveolar cells, whereas Elf-5 is dispensable for the induction of any $CD61^+$ lineages. Under this scenario, the increase in $CD61^+$ luminal progenitors observed in Elf-5 deficient glands is because of the requirement of Elf-5 for conversion of these cells into fully differentiated alveolar cells. Taken together, these results suggest that a transcriptional regulatory network involving STAT5A and Elf5 has a critical role in the establishment of alveolar cell identity.

In contrast to a specific role for STAT5A and Elf-5 in alveolar differentiation, GATA-3 function appears to be required at distinct stages of mammary gland development, first in the specification of the luminal cell fate and second during alveolar differentiation. It does not appear that GATA-3 lies downstream of STA5A/5B during alveolar differentiation given that GATA-3 levels are unchanged in STAT5A/5B-deficient glands (Yamaji

et al., 2009). It remains possible that GATA-3 may lie upstream of STAT5A/5B or Elf-5 during alveolar differentiation. In this regard, the analysis of transcription factors expressed in terminal end-buds and mature ductal epithelium, a pattern that mimicked GATA-3 expression, identified Elf-5 (Kouros-Mehr *et al*., 2006). Whether GATA-3 functions within the same cells and can directly regulate the expression of STAT5A or Elf-5 requires further experimentation. It has also been shown that the expression patterns of GATA-3 and Elf-5 within the luminal epithelial compartment are largely distinct. In the case of CD61 + cells, GATA-3 and Elf-5 have both been shown to be expressed in this population; however, it has been suggested that these transcription factors may not be coexpressed within the same $CD61^+$ epithelial cell (Oakes *et al*., 2008). Further experimental work will be required to dissect the functional relationship between GATA-3, STAT5A and Elf-5 in alveolar cell differentiation.

Bridging mammary gland development and breast cancer

Gene expression profiling has categorized breast cancers into several subtypes including luminal A, luminal B, HER2, basal and normal-like (Perou *et al*., 2000; Sorlie *et al*., 2003), which are correlated with distinct clinical outcomes (Sorlie *et al*., 2001; Sotiriou *et al*., 2003). These original subtypes will undoubtedly be revised as additional subtypes emerge, such as the putative 'claudin-low' subtype (Herschkowitz *et al*., 2007; Hennessy *et al*., 2009). An open question in the field of breast cancer is whether the different sub-types of breast cancer, which are defined by gene-expression profiling, reflect the transformation of specific stem cell or progenitor cell compartments responsible for development of the normal mammary gland (Prat and Perou, 2009). Using tissue from human breast reduction mammoplasties, cells were enriched for specific markers and were divided into one of four groups, including mammary stem-cell-enriched (MaSC), luminal progenitor, mature luminal and stromal populations (Lim *et al*., 2009). When the profiles of these four populations was compared with the previously defined breast cancer subtypes, it was found that the mammary stem-cell-enriched signature was most prominent in the 'claudin-low' and normal-like subtypes, the mature luminal signature was found highly represented in the luminal A, luminal B and HER2 subtypes, whereas the stromal signature was associated with the 'claudin-low' subtype. Perhaps the most surprising finding was the correlation between the luminal progenitor signature and the basal subtype (Lim *et al*., 2009). This study suggests that the differentiation hierarchy in the developing mammary gland might indeed have connections with the different subtypes of cancer that arise from this tissue. Moreover, emerging data suggests that those transcription factors that have critical roles in the specification of different mammary epithelial progenitors might also have important affects on cancer progression.

GATA-3 is a marker of the luminal subtype in breast cancer

As described above, GATA-3 specifies commitment of mammary stem cell/luminal progenitor cells to a mature luminal cell fate (Kouros-Mehr *et al*., 2006; Asselin-Labat *et al*., 2007). Thus, GATA-3 represents a strong candidate marker for the luminal breast cancer subtypes. Indeed, GATA-3 expression is strongly linked to ER-positive status, the luminal A subtype and good patient prognosis in numerous gene expression studies of primary human breast cancer (Mehra *et al*., 2005; Fang *et al*., 2009). Recent studies with the PyV mT transgenic mouse model (Guy *et al*., 1992; Lin *et al*., 2003) have examined potential roles played by GATA-3 in breast cancer development and progression. Comparison of gene expression signatures from early adenomas to those obtained from late carcinomas in the PyV mT model revealed that the expression of genes associated with luminal cell differentiation, including GATA-3, was dramatically impaired during this transition (Kouros-Mehr *et al*., 2008a). Conversely, when GATA-3 was ectopically expressed in PyV mT-expressing tumors, the GATA-3 expressing tumors possessed a well-differentiated phenotype that correlated with a dramatic reduction in their ability to metastasize when compared with parental PyV mT tumors (Kouros-Mehr *et al*., 2008a). However, when GATA-3 was deleted in well-differentiated tumors, it did not result in the progression to a more metastatic phenotype. Rather, breast cancer cells that lost GATA-3 underwent caspase-mediated cell death (Kouros-Mehr *et al*., 2008a). These observations have led to the hypothesis that, in the PyV mT transgenic model, GATA-3 promotes the formation of well-differentiated breast cancers that have a low metastatic potential. The eventual emergence of highly metastatic tumor cells in late carcinomas may reflect the expansion of a GATA-3 negative stem cell like population. The notion that GATA-3 expression is associated with good prognosis through its ability to promote differentiation in breast cancers may not fully explain how GATA-3 functions to suppress the metastatic properties of breast cancers. It has recently been demonstrated that GATA-3 overexpression in MDA-MB-231 breast cancer cells, which are a basal human breast cancer cell population, could impair mammary tumor growth and lung metastasis without promoting the differentiation of these cells (Dydensborg *et al*., 2009). Whether GATA-3 directly modulate the expression of genes involved in breast cancer metastasis, independently of its effects on breast cancer differentiation, will require further investigation.

Is STAT5 involved in mammary tumor progression?

Although a clear role for STAT5A in mammary development has been established, STAT5A function in breast cancer progression remains unclear. High levels of tyrosine-phosphorylated and nuclear-localized STAT5A can be detected in human breast cancer

samples (Cotarla *et al*., 2004). Furthermore, reduction of STAT5A dosage impairs mammary tumor induction in mice-expressing SV40T antigen (Ren *et al*., 2002). Other studies have suggested that signaling elements that act upstream of STAT5A may have an important role in mammary cancer development. For example, the loss of either Prolactin or ErbB4, which are known to be potent activators of STAT5 signaling, also have major effects on alveolar differentiation (Jones *et al*., 1999; Clevenger, 2003). Loss of Prolactin signaling impairs mammary tumorigenesis in mice expressing the PyV mT oncogene (Vomachka *et al*., 2000). However, because mammary gland development is generally impaired in Prolactin-deficient mice, it is unclear whether the reduction in mammary tumor formation reflects an indirect impact on the development of the mammary epithelial cells that are targeted by PyV mT oncogene.

Although ErbB4 is a potent activator of Stat5A function, mammary epithelial disruption of ErbB4 had little impact on the ability of ErbB2 to induce mammary tumors (Jackson-Fisher *et al*., 2006). In fact, a recent study has suggested that catalytically inactive somatic mutations in ErbB4 can be detected in sporadic human breast cancer (Tvorogov *et al*., 2009). Significantly these inactivating mutations in ErbB4 result in dramatic impairment in the capacity to activate Stat5A and induce differentiation of MDA-MB-468 cells (Tvorogov *et al*., 2009).These data would argue that tyrosine phosphorylation of STAT5A may actually impair breast cancer development by inducing differentiation.

One of the future challenges in elucidating the role of STAT5A in breast cancer will be discriminating between the direct effects of STAT5 family signaling on tumor development and progression and the effects of STAT5A's indirect role in the development of mammary epithelial cells that are then the target for transformation by activated oncogenes. Mouse models that permit the conditional and inducible loss of STAT5A/5B in the mammary gland, like those described by the Hennighausen group (Yamaji *et al*., 2009), will have an instrumental role in assessing the precise role of STAT5 in both initiation and maintenance of oncogene-induced mammary tumor progression.

A STAT3 dependent transcription regulatory network is involved in mammary gland involution and metastasis

Like the STAT5A and GATA-3 transcription factors, the STAT3 transcription factor has been implicated as a key regulator of both normal mammary gland development and mammary tumor progression (Pensa *et al*., 2009). The development of transgenic mice that carry conditional alleles of STAT3 has been critically important for understanding STAT3's role in mammary gland biology. Using different mammary epithelial Cre-expressing transgenic reporter strains, it has been demonstrated that mammary-epithelial disruption of STAT3 results in a dramatic impairment of mammary gland involution (Chapman *et al*., 1999; Humphreys *et al*., 2002). A survey of STAT3 dependent transcriptional targets has revealed that many are involved in acute inflammation and acute-phase regulation often associated with mammary gland involution. One such gene encodes the Oncostatin M receptor (OSMR) that, when activated, results in enhanced STAT3 tyrosine phosphorylation. STAT3 activation further enhances OSMR expression, thereby establishing a positive-feedback loop. Conversely, OSMR activation results in the dephosphorylation of STAT5 (Tiffen *et al*., 2008). Consistent with its importance in promoting mammary gland involution, OSMR-deficient mice still retained mammary epithelium even after 10 days of involution (Tiffen *et al*., 2008). The reciprocal regulation of STAT5A and STAT3 by OSMR suggests that these closely related STAT transcription factors have opposing roles in mammary gland involution.

Another important transcriptional target of STAT3 was C/EBPδ (Thangaraju *et al*., 2005; Zhang *et al*., 2007). Like STAT3, germ-line disruption of C/EBPδ resulted in a profound defect in mammary gland involution (Thangaraju *et al*., 2005). Thus, genetic and biochemical analyses suggest that a STAT3-dependent inflammatory cascade is essential for the normal involution of the mammary gland (Pensa *et al*., 2009). STAT3 acts to invoke a massive inflammatory response that is required for efficient clearance of apoptotic mammary epithelium during involution. In contrast to STAT5, which acts in a cell-autonomous manner to determine the luminal alveolar cell fate, STAT3 functions in non-cell autonomous manner to modulate the mammary microenvironment during mammary gland involution.

STAT3's role as the regulator of a pro-apoptotic program in normal mammary gland development is in apparent contradiction to its documented pro-survival functions in mammary tumor progression (Guo *et al*., 2006; Zhou *et al*., 2007). A potential explanation may lie in recent work demonstrating that the involuting mammary gland is a particularly good microenvironment for promoting mammary tumor metastasis (McDaniel *et al*., 2006; Schedin *et al*., 2007). Recently published data suggests that the pro-inflammatory functions of STAT3 during mammary gland involution may also have a critical role in mammary tumor progression. In this work, mammary epithelial disruption of STAT3 in transgenic mice expressing the ErbB2 oncogene provide compelling evidence that a STAT3-dependent pro-inflammatory program is required for the metastatic phase of tumor progression (Ranger *et al*., 2009). Although mammary tumor onset was not affected by ablation of STAT3, the resulting tumors rarely metastasized to the lung (Ranger *et al*., 2009). Remarkably, analyses of gene expression profile of these STAT3-deficient ErbB2 tumors revealed that the acute inflammatory response was dramatically impaired because of the reduced expression of C/EBPδ, OSMR, SAA1 and SAA2. The fact that this same STAT3-dependent transcriptional regulatory network was involved in regulating mammary gland involution (Pensa *et al*., 2009), suggests that STAT3 may use the same network in mammary tumor progression. Consistent

with this concept, mammary specific disruption of the STAT3 co-activator SRC-1 suppresses tumor metastasis without affecting the primary growth of PyV mT mammary tumors (Wang *et al*., 2009). Taken together, these observations argue that, as in normal mammary gland involution, STAT3's primary function during mammary tumor progression is to alter the immediate tumor microenvironment.

Future perspectives

One of the key issues stemming from a comparison of these transcription factor networks in normal mammary gland development is to what extent they are utilized during tumor progression. Emerging data suggests that distinct progenitor cells in the developing mammary gland, which are specified through the action of these transcriptional regulatory networks, are targeted by oncogenic initiating events which ultimately give rise to the different breast cancer subtypes (Lim *et al*., 2009; Prat and Perou, 2009). However, the tremendous heterogeneity that occurs within a given breast cancer suggests that these tumor cells are under considerable evolutionary pressures (Campbell and Polyak, 2007; Polyak, 2007a, 2007b). Accumulated genetic lesions within a breast cancer cell will have an impact on how these networks interact and will thus ultimately shape the tumor phenotype. In addition to epithelial cell-autonomous consequences, activation of these transcription networks can also have an impact on other cell types within the tumor microenvironment and thereby influence their capacity to metastasize (Ranger *et al*., 2009). The future elucidation of the complex relationship between the normal and pathological outputs of these networks may provide important therapeutic insights into the treatment of breast cancer.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Dr Josie Ursini-Siegel and Rob Annan for their critical reading of the manuscript. Work carried out in the laboratory of Dr WJ Muller was supported by a grant from the National Cancer Institute of Canada/Terry Fox Foundation Program Project Grant (#017003). PMS is a research scientist of the Canadian Cancer Society and WJM holds a Canada Research Chair in Molecular Oncology.

References

Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC *et al*. (2007). Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat Cell Biol* **9**: 201–209.

Asselin-Labat ML, Vaillant F, Shackleton M, Bouras T, Lindeman GJ, Visvader JE. (2008). Delineating the epithelial hierarchy in the mouse mammary gland. *Cold Spring Harb Symp Quant Biol* **73**: 469–478.

Brisken C. (2002). Hormonal control of alveolar development and its implications for breast carcinogenesis. *J Mammary Gland Biol Neoplasia* **7**: 39–48.

Campbell LL, Polyak K. (2007). Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle* **6**: 2332–2338.

Chapman RS, Lourenco PC, Tonner E, Flint DJ, Selbert S, Takeda K *et al*. (1999). Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3. *Genes Dev* **13**: 2604–2616.

Choi YS, Chakrabarti R, Escamilla-Hernandez R, Sinha S. (2009). Elf5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: failure of Stat5 activation and functional differentiation in the absence of Elf5. *Dev Biol* **329**: 227–241.

Clevenger CV. (2003). Role of prolactin/prolactin receptor signaling in human breast cancer. *Breast Dis* **18**: 75–86.

Cotarla I, Ren S, Zhang Y, Gehan E, Singh B, Furth PA. (2004). Stat5a is tyrosine phosphorylated and nuclear localized in a high proportion of human breast cancers. *Int J Cancer* **108**: 665–671.

Dydensborg AB, Rose AA, Wilson BJ, Grote D, Paquet M, Giguere V *et al*. (2009). GATA3 inhibits breast cancer growth and pulmonary breast cancer metastasis. *Oncogene* **28**: 2634–2642.

Eirew P, Stingl J, Raouf A, Turashvili G, Aparicio S, Emerman JT *et al*. (2008). A method for quantifying normal human mammary epithelial stem cells with *in vivo* regenerative ability. *Nat Med* **14**: 1384–1389.

Fang SH, Chen Y, Weigel RJ. (2009). GATA-3 as a marker of hormone response in breast cancer. *J Surg Res* **157**: 290–295.

Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M *et al*. (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **1**: 555–567.

Grimm SL, Rosen JM. (2003). The role of C/EBPbeta in mammary gland development and breast cancer. *J Mammary Gland Biol Neoplasia* **8**: 191–204.

Guo W, Pylayeva Y, Pepe A, Yoshioka T, Muller WJ, Inghirami G *et al*. (2006). Beta 4 integrin amplifies ErbB2 signaling to promote mammary tumorigenesis. *Cell* **126**: 489–502.

Guy CT, Cardiff RD, Muller WJ. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol* **12**: 954–961.

Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS *et al*. (2009). Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res* **69**: 4116–4124.

Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z *et al*. (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **8**: R76.

Humphreys RC, Bierie B, Zhao L, Raz R, Levy D, Hennighausen L. (2002). Deletion of Stat3 blocks mammary gland involution and extends functional competence of the secretory epithelium in the absence of lactogenic stimuli. *Endocrinology* **143**: 3641–3650.

Jackson-Fisher AJ, Bellinger G, Shum E, Duong JK, Perkins AS, Gassmann M *et al*. (2006). Formation of Neu/ErbB2-induced mammary tumors is unaffected by loss of ErbB4. *Oncogene* **25**: 5664–5672.

Jones FE, Welte T, Fu XY, Stern DF. (1999). ErbB4 signaling in the mammary gland is required for lobuloalveolar development and Stat5 activation during lactation. *J Cell Biol* **147**: 77–88.

Kouros-Mehr H, Bechis SK, Slorach EM, Littlepage LE, Egeblad M, Ewald AJ *et al*. (2008a). GATA-3 links tumor differentiation and dissemination in a luminal breast cancer model. *Cancer Cell* **13**: 141–152.

Kouros-Mehr H, Kim JW, Bechis SK, Werb Z. (2008b). GATA-3 and the regulation of the mammary luminal cell fate. *Curr Opin Cell Biol* **20**: 164–170.

Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. (2006). GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* **127**: 1041–1055.

Lamarca HL, Visbal AP, Creighton CJ, Liu H, Zhang Y, Behbod F *et al*. (2010). C/EBPbeta Regulates stem cell activity and specifies luminal cell fate in the mammary gland. *Stem Cells* (e-pub ahead of print).

Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH *et al*. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15**: 907–913.

Lin EY, Jones JG, Li P, Zhu L, Whitney KD, Muller WJ *et al*. (2003). Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am J Pathol* **163**: 2113–2126.

Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A, Hennighausen L. (1997). Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes Dev* **11**: 179–186.

McDaniel SM, Rumer KK, Biroc SL, Metz RP, Singh M, Porter W *et al*. (2006). Remodeling of the mammary microenvironment after lactation promotes breast tumor cell metastasis. *Am J Pathol* **168**: 608–620.

Mehra R, Varambally S, Ding L, Shen R, Sabel MS, Ghosh D *et al*. (2005). Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. *Cancer Res* **65**: 11259–11264.

Miyoshi K, Shillingford JM, Smith GH, Grimm SL, Wagner KU, Oka T *et al*. (2001). Signal transducer and activator of transcription (Stat) 5 controls the proliferation and differentiation of mammary alveolar epithelium. *J Cell Biol* **155**: 531–542.

Molyneux G, Regan J, Smalley MJ. (2007). Mammary stem cells and breast cancer. *Cell Mol Life Sci* **64**: 3248–3260.

Naylor MJ, Ormandy CJ. (2007). Gata-3 and mammary cell fate. *Breast Cancer Res* **9**: 302.

Oakes SR, Naylor MJ, Asselin-Labat ML, Blazek KD, Gardiner-Garden M, Hilton HN *et al*. (2008). The Ets transcription factor Elf5 specifies mammary alveolar cell fate. *Genes Dev* **22**: 581–586.

Pensa S, Watson CJ, Poli V. (2009). Stat3 and the inflammation/acute phase response in involution and breast cancer. *J Mammary Gland Biol Neoplasia* **14**: 121–129.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA *et al*. (2000). Molecular portraits of human breast tumours. *Nature* **406**: 747–752.

Polyak K. (2007a). Breast cancer stem cells: a case of mistaken identity? *Stem Cell Rev* **3**: 107–109.

Polyak K. (2007b). Breast cancer: origins and evolution. *J Clin Invest* **117**: 3155–3163.

Prat A, Perou CM. (2009). Mammary development meets cancer genomics. *Nat Med* **15**: 842–844.

Ranger JJ, Levy DE, Shahalizadeh S, Hallett M, Muller WJ. (2009). Identification of a Stat3-dependent transcription regulatory network involved in metastatic progression. *Cancer Res* **69**: 6823–6830.

Ren S, Cai HR, Li M, Furth PA. (2002). Loss of Stat5a delays mammary cancer progression in a mouse model. *Oncogene* **21**: 4335–4339.

Schedin P, O'Brien J, Rudolph M, Stein T, Borges V. (2007). Microenvironment of the involuting mammary gland mediates mammary cancer progression. *J Mammary Gland Biol Neoplasia* **12**: 71–82.

Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML *et al*. (2006). Generation of a functional mammary gland from a single stem cell. *Nature* **439**: 84–88.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H *et al*. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* **98**: 10869–10874.

Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A *et al*. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* **100**: 8418–8423.

Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A *et al*. (2003). Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci USA* **100**: 10393–10398.

Sternlicht MD, Kouros-Mehr H, Lu P, Werb Z. (2006). Hormonal and local control of mammary branching morphogenesis. *Differentiation* **74**: 365–381.

Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D *et al*. (2006a). Purification and unique properties of mammary epithelial stem cells. *Nature* **439**: 993–997.

Stingl J, Raouf A, Eirew P, Eaves CJ. (2006b). Deciphering the mammary epithelial cell hierarchy. *Cell Cycle* **5**: 1519–1522.

Thangaraju M, Rudelius M, Bierie B, Raffeld M, Sharan S, Hennighausen L *et al*. (2005). C/EBPdelta is a crucial regulator of pro-apoptotic gene expression during mammary gland involution. *Development* **132**: 4675–4685.

Tiffen PG, Omidvar N, Marquez-Almuina N, Croston D, Watson CJ, Clarkson RW. (2008). A dual role for oncostatin M signaling in the differentiation and death of mammary epithelial cells *in vivo*. *Mol Endocrinol* **22**: 2677–2688.

Tvorogov D, Sundvall M, Kurppa K, Hollmen M, Repo S, Johnson MS *et al*. (2009). Somatic mutations of ErbB4: selective loss-of-function phenotype affecting signal transduction pathways in cancer. *J Biol Chem* **284**: 5582–5591.

Visvader JE. (2009). Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. *Genes Dev* **23**: 2563–2577.

Vomachka AJ, Pratt SL, Lockefeer JA, Horseman ND. (2000). Prolactin gene-disruption arrests mammary gland development and retards T-antigen-induced tumor growth. *Oncogene* **19**: 1077–1084.

Wang S, Yuan Y, Liao L, Kuang SQ, Tien JC, O'Malley BW *et al*. (2009). Disruption of the SRC-1 gene in mice suppresses breast cancer metastasis without affecting primary tumor formation. *Proc Natl Acad Sci USA* **106**: 151–156.

Yamaji D, Na R, Feuermann Y, Pechhold S, Chen W, Robinson GW *et al*. (2009). Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. *Genes Dev* **23**: 2382–2387.

Zhang Y, Sif S, DeWille J. (2007). The mouse C/EBPdelta gene promoter is regulated by STAT3 and Sp1 transcriptional activators, chromatin remodeling and c-Myc repression. *J Cell Biochem* **102**: 1256–1270.

Zhou J, Wulfkuhle J, Zhang H, Gu P, Yang Y, Deng J *et al*. (2007). Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance. *Proc Natl Acad Sci USA* **104**: 16158–16163.