Journal of Mammary Gland Biology and Neoplasia (2019) 24:213–230
https://doi.org/10.1007/s10911-019-09430-6

Emerging Role of SOX Proteins in Breast Cancer Development and Maintenance

Gaurav A. Mehta[1,2,3] · Pooja Khanna[1,2,3] · Michael L. Gatza[1,2,3]

Received: 8 February 2019 / Accepted: 21 April 2019 / Published online: 9 May 2019

Abstract

The *SOX* genes encode a family of more than 20 transcription factors that are critical regulators of embryogenesis and developmental processes and, when aberrantly expressed, have been shown to contribute to tumor development and progression in both an oncogenic and tumor suppressive role. Increasing evidence demonstrates that the SOX proteins play essential roles in multiple cellular processes that mediate or contribute to oncogenic transformation and tumor progression. In the context of breast cancer, SOX proteins function both as oncogenes and tumor suppressors and have been shown to be associated with tumor stage and grade and poor prognosis. Experimental evidence demonstrates that a subset of SOX proteins regulate critical aspects of breast cancer biology including cancer stemness and multiple signaling pathways leading to altered cell proliferation, survival, and tumor development; EMT, cell migration and metastasis; as well as other tumor associated characteristics. This review will summarize the role of SOX family members as important mediators of tumorigenesis in breast cancer, with an emphasis on the triple negative or basal-like subtype of breast cancer, as well as examine the therapeutic potential of these genes and their downstream targets.

Keywords Sox · Breast cancer · Oncogene · Cancer stem cells · EMT · Signaling

Introduction

Breast cancer is the most commonly diagnosed form of cancer and the second leading cause of cancer related deaths among women in the United States [1]. Despite significant advances in diagnostic and treatment strategies, approximately 270,000 new cases are diagnosed and 40,000 deaths reported annually in the United States [1]. The difficulties in detecting and developing effective therapeutic strategies are due, in part, to the underlying genetic and molecular heterogeneity that give rise to clinical variability that is characteristic of this disease [2–5]. Clinically, breast cancers are classified on the basis of expression of the estrogen receptor (ER), progesterone receptor (PR) and HER2 oncogene as ER positive, HER2 positive or triple negative (*i.e.* negative for all three markers) breast tumors (TNBC). Seminal studies by Perou and colleagues, as well as a multitude of other reports, have demonstrated additional molecular heterogeneity within breast cancer by identifying molecular subtypes based on gene expression and genomic or proteomic profiling [4, 6–10]. The PAM50 subtypes, the most prominent of these classification strategies, identified five molecular subtypes of breast cancer based on gene expression profiling: basal-like, HER2 enriched (HER2E), luminal A, luminal B, and normal-like. These subtypes, in addition to the more recently identified claudin-low subtype, differ significantly with respect to underlying biology, as well as in terms of incidence, response to therapy and clinical outcomes [2, 6, 8]. While the luminal subtypes of breast cancer are predominantly estrogen receptor positive (ER+), basal-like breast cancers, which account for 70-80% of TNBCs, are largely negative for ER, PR and HER2 expression and are characterized by high rates of cell proliferation, aggressive clinical behavior and have the worst prognosis. Basal-like tumors are predominant in African-American women as well as younger women and have the shortest overall survival rate and highest incidence of relapse [2, 11, 12]. Evidence suggests that at the molecular level, these tumors are defined by a unique set of genetic alterations leading to altered cellular

✉ Michael L. Gatza
michael.gatza@cinj.rutgers.edu

[1] Rutgers Cancer Institute of New Jersey, 195 Little Albany Street, CINJ 4558, New Brunswick, NJ 08903, USA

[2] Department of Radiation Oncology, Robert Wood Johnson Medical School, New Brunswick, NJ, USA

[3] Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

signaling and, as such, are highly variable in terms of their chemotherapy sensitivity [3, 9, 13, 14]. Further, gene expression profiling studies from The Cancer Genome Atlas pan-cancer project clearly demonstrate that basal-like breast cancers are significantly different from non-basal-like tumors, which is consistent with previous studies that suggest these two classes of tumors may arise from distinct cellular origins and/or may evolve to mimic specific cellular states [15–17]. Thus, TNBC constitute a unique disease entity that poses a significant clinical challenge as these cancers do not respond to hormonal therapy and are largely refractive to available targeted agents. As such, cytotoxic chemotherapy, despite its limited efficacy and toxic side effects, remains the current standard-of-care treatment for these patients. Due to the complex and heterogeneous nature of triple-negative or basal-like breast cancers as well as the lack of effective therapies, there is an urgent need to better understand the molecular and genetic mechanisms altered in these tumors in order to enable the development of novel and rational therapeutic strategies based on the underlying biology of the disease. Consistent with these ideas, a number of recent studies have demonstrated that multiple members of the SOX transcription factor family are overexpressed and activated in TNBC or basal-like tumors and emerging data provide evidence that this family of proteins play an essential role in tumor development and progression. In this review, we will summarize the functions of each SOX family member and the role it plays in the development and maintenance of breast cancer, with an emphasis on the basal-like subtype of breast cancer.

Overview of SOX Gene and Protein Classification, Structure and Function

The proteins of the SRY-related HMG-box (SOX) family were first identified based on their sequence similarity with the HMG (high mobility group) DNA-binding domain of the *SRY* gene [18–20]. The HMG-box is a 79-amino acid domain that allows for interaction of the SOX proteins with the A/TA/TCAAA/TG motif in the minor groove of the DNA [21, 22]. Since the discovery of the first SOX proteins in the 1990's, twenty-one SOX family members with overlapping and divergent functions have been identified in the vertebrate genome and shown to affect various cellular functions, often in a context and tissue-specific manner [23–49]. These proteins have been classified into eight groups based on HMG domain sequence, protein structure and evolutionary relationships as illustrated in Fig. 1. These groups are: A, B (comprised of B1 and B2 subgroups), C, D, E, F, G and H [18, 50, 51]. In humans, members of each of these groups are: **SOXA:** *SRY*; **SOXB:** *SOX1*, *SOX2*, *SOX3*, *SOX14* and *SOX21*; **SOXC:** *SOX4*, *SOX11* and *SOX12*; **SOXD:** *SOX5*, *SOX6* and *SOX13*; **SOX E:** *SOX8*, *SOX9* and *SOX10*; **SOXF:** *SOX7*, *SOX17* and *SOX18*; **SOXG:** *SOX15* and *SOX20*; and **SOXH:** *SOX30*.

Although each of these proteins recognize the same consensus sequence, no common set of target genes have been identified and no single common biological role has been attributed to the activity of these proteins. While several studies suggest that some functional overlap may exist between various SOX family members, the specific mechanisms by which these proteins affect cellular activities has been found to be unique to each SOX class, and in some cases, each SOX protein. This is not surprising given that the amino acid sequence outside the HMG box domain, which determines the transcriptional specificity of the SOX proteins through interaction with various cofactors or transcriptional regulators, show little to no overlap between groups [18–20]. Interestingly, evidence strongly suggests that many SOX proteins may function in a tissue-specific and context-specific manner, which may further complicate our understanding of the impact this transcription factor family has on breast cancer biology [18, 52]. In general, strong evidence demonstrates that SOX proteins are essential for embryonic and mammary gland development. These data further highlight an important role - both oncogenic and tumor suppressive - for a subset of SOX transcription factor family members in regulating critical aspects of breast cancer biology including multiple facets of breast cancer genesis, progression and therapeutic response [19, 20, 53].

Clinical Relationship Between SOX Proteins and Breast Cancer

Aberrant SOX gene and protein expression in human breast tumors has been observed in multiple studies and emerging experimental data suggest that altered activation of this gene family may contribute to key aspects of breast cancer genesis and progression. We have illustrated the mRNA expression patterns of SOX family genes by PAM50 subtype in 1,052 primary tumors and 94 adjacent normal samples from The Cancer Genome Atlas (TCGA) project (Fig. 2). These analyses clearly show altered expression of several *SOX* genes relative to adjacent normal breast tissue and within the context of the PAM50 molecular subtypes. Interestingly, studies have suggested both an oncogenic and tumor suppressive role of specific SOX family proteins and expression of these genes or proteins often corresponds with clinical characteristics including prognosis and metastasis.

Clinically, increased *SOX4* expression has been observed in multiple tumor types including breast, prostate, bladder, hepatocellular carcinoma, medullobastoma and small cell lung carcinoma [54–60]. In breast cancer, both *SOX4* mRNA and protein levels have been found to be highly up-regulated compared to adjacent normal tissue [55] (Fig. 2).

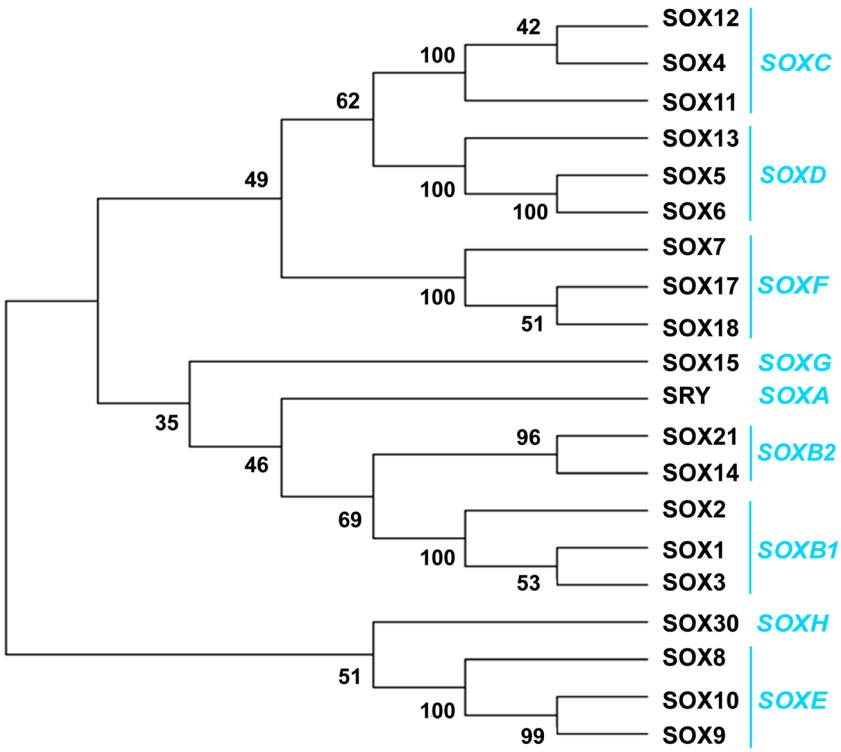

Fig. 1 Groups and phylogenetic tree of human SOX proteins. A rooted phylogenetic neighbor-joining tree for the human SOX full-length proteins was performed based on conserved amino acid sequences during evolution and divergence. To determine the robustness of the phylogeny relationship, 1000 bootstrap replicates were carried out. Each (%) bootstrap value is shown at the branch points

Additional studies have shown increased *SOX4* expression as well as increased DNA amplification frequency in human breast tumors and demonstrated that this observation is largely specific to basal-like or TNBC tumors [54]. Consistent with these data, immunohistochemical (IHC), proteomic and transcriptomic-based analyses have reported that *SOX4* expression and/or activity corresponds with a poor overall prognosis for breast cancer patients and, in particular, for basal-like or TNBC patients [55, 61, 62]. These studies collectively indicate that *SOX4* expression corresponds with increased tumor aggressiveness.

Additional SOXC proteins have also been shown to be aberrantly overexpressed in TNBC tumors and associated with poor survival [32, 34]. In particular, *SOX11*, an oncogene with increased expression in basal-like tumors (Fig. 2), was identified in a large-scale genetic screen as an essential transcription factor required for proliferation and metastatic phenotypes in basal-like tumors. However, consistent with observed patterns of expression in breast tumors, *SOX11* was not found to be essential in other breast cancer subtypes [32]. Not surprisingly, *SOX11* has been identified as a marker of poor survival in basal-like tumors. Finally, *SOX12,* which is

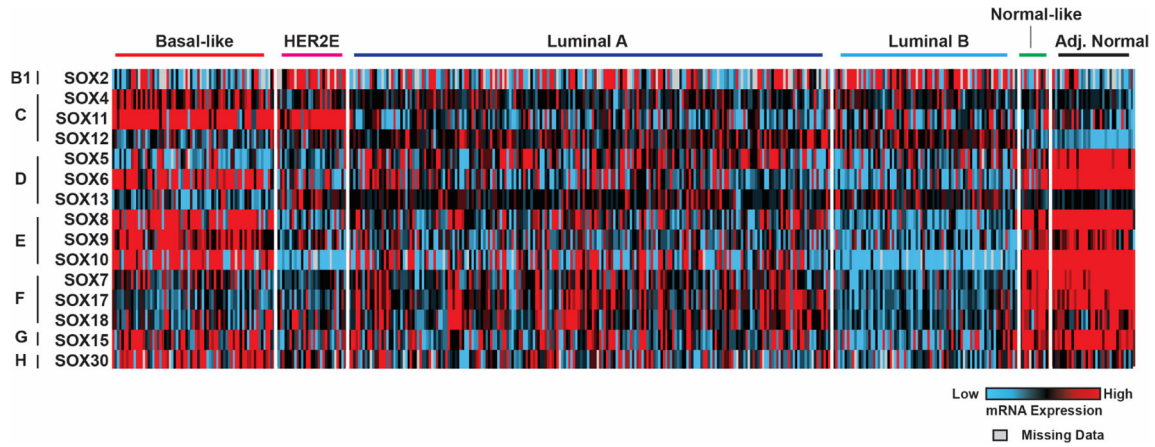

Fig. 2 Analysis of SOX family member mRNA expression by breast cancer subtype. Patterns of SOX gene expression were determined for 1,052 human breast tumors and 94 adjacent normal samples from the TCGA dataset; red indicates high mRNA expression and blue depicts low mRNA levels. Samples are organized by PAM50 molecular subtype: Basal-like ($n = 185$), HER2 Enriched (HER2E; $n = 79$), Luminal A (LumA; $n = 545$), Luminal B (LumB; $n = 210$) and Normal-like ($n = 33$) tumors. SRY, SOX1, SOX3, SOX14, SOX20 and SOX21 were excluded from this analysis due to missing or insufficient data (expression values present in >80% of samples)

the least studied member of the SOXC family has been shown to be upregulated in breast cancer patient samples [34], although it does not appear to be uniformly expressed in any specific subtype as illustrated in Fig. 2.

Beyond the SOXC family members, SOXE proteins have also been shown to be consistently overexpressed in basal-like tumors (Fig. 2) [33, 44, 63]. The best studied of these proteins, SOX10, was found to be enriched in breast cancer patient samples, particularly basal-like and TNBCs, as well as metastatic TNBCs and secretory carcinomas [23–25, 27, 28, 30, 37, 42–44]. These findings appear to be somewhat controversial, as other studies have reported lower nuclear *SOX10* expression in TNBCs compared to ER+/luminal and Her2+ tumors [36]. Likewise, analyses of tumor versus peripheral normal tissue showed no differences in *SOX10* expression in early stage (pT1 and pT2 or pN0 and pN1) tumors [64]. However, the results of these apparently conflicting studies may be more consistent than expected since percentages of SOX10 positive TNBC and luminal/HER2+ tumors were more comparable if strong and mildly positive cases were considered as a single class in the former study and if it is appreciated that samples analyzed in the latter study were early stage tumors [36, 64]. It is clear however, that additional analyses are needed in early and late stage primary samples as well as metastatic tissue to determine the distribution of SOX10 in these tumors.

SOX8 is the least studied member of the SOXE family in mammary tumorigenesis and its role in breast cancer biology is poorly understood. However, Dong *et al*. have demonstrated that the *SOX8* gene is amplified in about 1.6% of breast cancer patient samples from their dataset and that *SOX8* DNA copy number status was indicative of poor overall survival in these patients [63]. Finally, while *SOX9* is the most well characterized member of the SOXE family, and is predominantly overexpressed in basal-like tumors (Fig. 2), no evidence currently exists demonstrating its prognostic capacity. However, *SOX9* is oncogenic and, as outlined in detail in subsequent sections, has been shown to be essential for lineage commitment, differentiation and EMT during embryonic development as well as being crucial for oncogenesis through regulation of cancer stem cell population in breast tumors [65].

The SOXB family of proteins appears to be somewhat dichotomous with respect to their role in breast cancer genesis. Recent studies have demonstrated that *SOX2* is overexpressed in early stage breast carcinoma and is positively correlated with tumor size [66]. *SOX2* was found to be more frequently expressed in tumors with basal-like and TNBC phenotypes compared to other subtypes, has been shown to promote increased cell proliferation and metastasis, and is associated with shorter overall and disease-free survival [45, 47, 67]. Collectively, these data suggest that *SOX2*, like SOXC and SOXE family members, functions as an oncogene and is a critical determinant of survival in breast cancer patients. Conversely, *SOX1*, like SOXF family members *SOX7* and *SOX17* discussed below, appears to be tumor-suppressive and is frequently down regulated in breast cancer cell lines and patient tissue samples [31].

Members of the SOXF family of proteins demonstrate opposing roles in breast cancer with *SOX18* acting as an oncogene whereas *SOX7* and *SOX17* function as tumor suppressors [38, 40, 46]. IHC analysis of clinical samples from 122 Invasive Ductal Breast Carcinoma cases identified a significant positive correlation between SOX18 expression and malignancy grade [41]. Likewise, *SOX18* expression was strongly correlated with HER2 status and increased expression was observed in HER2+ cell lines compared to TNBC or normal breast epithelial cell lines [41]. In contrast to *SOX18*, expression of *SOX7* and *SOX17* is significantly decreased in breast cancer cell lines and tumor samples due to promoter hypermethylation and through regulation by microRNAs [38, 40, 46, 68–71]. Importantly, data demonstrate that higher *SOX7* and *SOX17* expression corresponds with increased metastasis free survival [38, 40, 72, 73]. In agreement with this, our own analysis of the TCGA dataset (Fig. 2) indicate that *SOX7* and *SOX17* are largely expressed at lower levels whereas *SOX18* is expressed at higher levels in human tumors relative to adjacent normal tissue. However, the relationship between the expression of these genes and clinical characteristics remains to be fully elucidated and additional studies will be necessary to fully establish the association between SOXF gene and/or protein expression and clinical characteristics in breast cancer.

The Impact of SOX Proteins on Cancer Stem Cells

The association between the signaling necessary for embryogenesis as well as mammary gland development and the reactivation or aberrant activation of these networks in breast cancer and, in particular, TNBC or basal-like breast cancer genesis has been extensively investigated [74–78]. These studies suggest that many solid tumors, including breast cancers, arise from cancer stem cells (CSCs) which are analogous to blastocyst derived embryonic stem cells (ESC), have the capacity to self-renew and give rise to heterogeneous, more differentiated cells with less proliferative capacity [74, 79]. Although many findings in the CSC field remain to be fully elucidated, a number of studies have proposed that these cells contribute to the phenotypic and functional heterogeneity observed in different cancer types [80, 81]. Moreover, studies have also shown that CSCs play a critical role in the therapeutic resistance and relapse observed in breast cancer [29, 77, 82] with several different effector molecules including transcription factors, chromatin remodelers and microRNAs (miRNAs) implicated in determining the fate of these cells in cancer [83].

SOX proteins are evolutionary conserved transcription factors and are amongst the earliest class of transcription factors to be expressed during embryogenesis and development [50, 83, 84]. Increasing evidence supports the role of these factors as critical regulators of stem cell fate with several members of the SOX family including SOX2 [85], SOX4 [86], SOX9 [87], SOX10 [88, 89], and SOX11 [90] contributing to the regulation of pluripotency in ESCs. As would be expected, a growing body of evidence strongly implicates the contribution of SOX protein to CSC phenotypes observed in TNBC or basal-like breast tumors.

SOX2 has been shown to be expressed early during development and is essential in the generation and maintenance of the pluripotent stem cell population [85, 91, 92]. Deletion of this essential gene *in vivo* results in embryonic lethality and a failure to generate pluripotent stem cells during development [92, 93]. Moreover, SOX2 in combination with OCT4 and MYC, has been shown to be essential for the formation of induced pluripotent stem cells (iPSCs) [92–94]. Consistent with its role in maintaining the stemness of embryonic stem cells, *SOX2* expression is altered in several tumor types with varying degrees of differentiation [85, 95–97]. Evidence suggests that this gene plays a role in defining the characteristics of the less-differentiated 'stem-cell' phenotype associated with basal-like breast tumor [45, 47]. Notably, Leis *et al.* demonstrated that *SOX2* expression was induced in tumor spheres from natural breast tumor cultures and breast carcinoma cell lines [45]. Overexpression and knockdown studies further showed that *SOX2* was sufficient to induce tumor sphere formation and tumor initiation *in vivo* indicating that SOX2 plays an important role in maintaining the cancer stem cell population [45].

More recently, VEGF was found to promote the breast cancer stem cell population by upregulating *MYC* and *SOX2* expression, leading to the induction of tumor sphere formation and aldehyde dehydrogenase activity in TNBC tumors and cell lines [79]. Moreover, inhibition of *SOX2* expression by TRPS1 (Transcriptional Repressor GATA Binding Protein 1) resulted in reduced mammosphere formation *in vitro* and decreased tumor burden in cell line-derived xenograft mouse models. These data indicated that inhibiting SOX2 activity resulted in suppression of cancer stemness and tumorigenic capacity [98].

As a key regulator of oncogenesis, SOX2 has been shown to activate the expression of, and be regulated by, a number of microRNAs (miRs). Recently, Deng *et al*. showed that breast cancer cell lines transfected with miR-378 acquire stem cell properties with increased cell survival and colony formation capabilities [99]. In this study, the authors clearly showed that overexpression of miR-378 resulted in increased *SOX2* expression through suppression of vimentin (*VIM)*, which has been shown to inhibit *SOX2* expression in breast cancer cells [99]. Finally, studies by Zhang and colleagues demonstrated that ERα signaling can also regulate breast cancer stem cells by inhibiting the expression of miR-140 which was shown to target *SOX2* [100]. Consistent with these findings, SOX2 was shown to promote tamoxifen resistance in breast cancer cells [82]. In this study, Piva and colleagues demonstrated that tamoxifen-resistant cells had higher levels of *SOX2* expression and increased stem cell characteristics. The investigators further confirmed that overexpression of *SOX2* in MCF7 cells was sufficient to promote tamoxifen resistance while shRNA-mediated silencing of *SOX2* expression increased sensitivity [82].

Finally, it has been well documented that obesity in breast cancer is associated with a more aggressive phenotype and increased breast cancer mortality [101]. In a recent study, Picon-Ruiz *et al.* demonstrated that the interaction between cancer cells and adipocytes resulted in increased expression of pro-inflammatory cytokines leading to up-regulation of oncogenic signaling in breast cancer. Specifically, the authors demonstrated that cytokines expressed from adipocytes resulted in activation of *SRC* in cancer cells, which in turn led to increased expression of stem cell factors *SOX2*, *NANOG*, and *MYC*. Importantly, SOX2 induction of miR-302b was found to further stimulate *MYC* and *SOX2* expression which potentiated stem-like characteristics of these cells and contributed to accelerated tumor growth and progression [101].

Comparable to *SOX2*, *SOX4* is essential for development as *SOX4* knockout mice show embryonic lethality at E14 [49]. In addition, *SOX4* expression is significantly increased in normal mammary stem cells isolated from cultured mammospheres suggesting that *SOX4* is involved in maintaining the CSC population in breast cancer [86]. Consistent with these findings, a number of studies have begun to provide insight into the mechanisms by which SOX4 contributes to the CSC phenotype. It was recently reported that *SOX4* overexpression in MCF10A cells led to an increase in the $CD44^{hi}CD24^{lo}$ population of CSCs [62]. While the exact mechanisms by which SOX4 mediated changes in this cell population remains unclear, it was determined that *SOX4* overexpression led to anchorage independent cell growth. Moreover, investigators determined that overexpression of *SOX4* in combination with *Ras* was sufficient to induce tumor growth in a xenograft mouse model indicating that *SOX4* was essential for tumor initiation [62]. Further evidence demonstrated that *SOX4* expression was increased when MCF7 cells were cultured in 3D collagen scaffolds and this increase was concordant with the enrichment of stem cells and pro-angiogenic factors [102, 103].

Interestingly, *SOX4* is a direct target of the TGFβ pathway which has been shown to increase the stem-like properties of TNBC cells following chemotherapy, thereby contributing to drug resistance and relapse [104]. Consistent with this premise, it was recently reported that in glioma initiating cells (GICs) *SOX4* has been shown to mediate and maintain

stemness of these cells through the TGFβ-SOX4-SOX2 signaling axis [105]. Although SOX4 has been shown to regulate expression of both *SOX2* and *OCT4*, it remains to be determined if a similar mechanism is involved in maintaining the stemness of breast cancer cells.

Given its central role in embryogenesis, *SOX9* has been proposed to function as a stem cell factor with important roles in maintaining the stem cell population during embryogenesis and in adult tissues [19, 106, 107]. Consistent with this premise, *SOX9* has been shown to function in lineage commitment and in the maintenance of stem cell populations in the hair follicles of the adult skin as well as neural crest stem cells [19, 106, 107]. In the context of breast cancer, *SOX9* nuclear expression was found to be significantly enriched in TNBC tumors compared to ER+ and HER2+ breast cancers [36]. Notably, increased *SOX9* expression was associated with upregulation of the $CD44^{hi}CD24^{low}$ cancer stem cell phenotype as well as poor prognosis [33, 65].

A number of recent studies have begun to investigate the mechanisms by which *SOX9* mediates the CSC phenotype and to determine the impact of *SOX9* expression and *SOX9*-mediated stemness on tumor development. Guo and colleagues recently demonstrated that exogenous overexpression of *Slug* and *SOX9* was sufficient to convert differentiated luminal cells into mammary stem cells (MaSCs) with long term mammary gland-reconstituting ability [87]. This study demonstrated that expression of *SOX9* promoted the tumorigenic and metastatic seeding abilities of human breast cancer cells, indicating that *SOX9* could confer stem cell-like properties upon tumor cells [87]. More recently, the *SOX2-SOX9* signaling axis has been shown to regulate the breast cancer stem cell content and resistance to endocrine therapy. *SOX2* was shown to regulate the expression of *SOX9* and CRISPR/Cas mediated *SOX9* silencing impaired stem cell self-renewal and abrogated tamoxifen resistant breast tumor growth [108]. Interestingly, lineage tracing experiments in mice demonstrated that *SOX9* expression distinguishes the mammary ER+ and ER- luminal stem cell populations and predominantly directs the development and maintenance of ER- luminal cells [109].

Consistent with these findings, SOX9 has been shown to regulate the expression of *FXYD3*, an estrogen inducible gene which is a critical player in the regulation of ER+ breast cancer CSC function [110]. FXYD3 has been shown to interact with SRC and ERα to form an activated complex and mediate non-genomic estrogen signaling. A number of studies have now demonstrated that the SOX9/FXYD3/SRC axis is required for maintaining the CSC population which promotes endocrine resistance in ER+ breast cancer [110, 111]. Finally, SOX9 has been shown to promote the metastatic phenotype in response to mTOR inhibition by transcriptional upregulation of the key mTOR pathway mediators and stem-cell signatures in breast tumor cells [112].

Similar to *SOX2* in ER+ breast cancer, the function and activity of *SOX9* in breast cancer is also regulated post-transcriptionally by microRNAs. MiR-140, which has been shown to be activated by ERα and mediate *SOX2* expression [100], was shown to regulate basal CSC self-renewal and tumor formation *in vivo* primarily through the miR-140/ALDH1/SOX9 axis [113]. Restoration of miR-140 levels, either genetically or pharmacologically, by adding the epigenetic modulator sulforaphane to cell line media or to mouse diet, decreased *SOX9* and *ALDH1* levels *in vitro* and reduced tumor growth *in vivo* [113].

Similar to other SOXE family members, *SOX10* expression has also been shown to correlate with cancer stem cell signatures and phenotypes. This association has been confirmed by deletion and overexpression studies using both *in vitro* cell line-based studies and *in vivo* murine breast cancer models [88, 89, 114, 115]. Recent studies have demonstrated that *SOX10* is a marker of TNBCs [44, 116, 117]. In TNBC tumors, ectopic *SOX10* expression resulted in upregulation of nestin leading to an increase in $CD44^{hi}CD24^{lo}$ cells and mammosphere formation [30]. The observation that SOX10 positive cells exhibit neural crest features [88], further implicates this gene as playing a role in maintaining the cancer stem cell phenotype and suggesting that increased *SOX10* expression may be associated with poor survival, relapse and drug resistance in these tumors.

Finally, a single study has assessed the impact of *SOX11* on CSCs in breast tumors [90]. Like *SOX4*, overexpression of the SOXC family member *SOX11* enhanced the stem cell phenotype. Specifically, it was demonstrated that overexpression of *SOX11* in mammary epithelial cells resulted in up-regulation of the CSC marker, aldehyde dehydrogenase as well as an increase in the percentage of $CD44^{hi}CD24^{lo}$ expressing cancer stem cells. This study subsequently demonstrated that *SOX11* overexpression was associated with increased mammosphere formation [90].

Overall, it has been well documented that SOX proteins are essential for embryogenesis and development and emerging data clearly demonstrate the role of these transcription factors in promoting the cancer stem cell phenotype in breast cancer. However, significant challenges remain including elucidating the mechanisms by which *SOX2*, *SOX4*, *SOX9*, *SOX10* and *SOX11* mediate breast cancer stemness, determining the impact these SOX proteins may have on mammary stem cells and/or breast cancer stem cells, and investigating the concordance between mechanisms by which SOX proteins regulate the cancer stem cell phenotype in breast cancer and cancers in other tissue types. It will be essential to determine whether these mechanisms are specific to each protein or molecular/clinical subtype of breast cancer and to demonstrate to what degree tumor development, progression and therapeutic response are mediated through SOX protein activation of these processes.

Regulation of Cellular Signaling, Cell Proliferation and Tumor Growth

As outlined above, the SOX proteins have divergent functions and in terms of tumor development can act both in an oncogenic or tumor suppressive role. A number of studies have demonstrated that overexpression of members of the SOX protein family can mediate oncogenic transformation in breast cancer through modulation of cellular signaling pathways that lead to increased cell proliferation and survival. While numerous signaling pathways play an important role in breast cancer development, dysregulation of the TGFβ and Wnt/β-catenin signaling pathways have been implicated as predominant mechanisms by which SOX-family proteins mediate cellular transformation, although ample evidence suggests that additional tumorigenic signaling pathways are regulated by these proteins.

In terms of breast cancer, *SOX4* is perhaps the best studied of the oncogenic SOX family members. *SOX4* has been shown to be oncogenic as overexpression of *SOX4* in combination with *Ras* can lead to transformation of mammary epithelial cells *in vitro* [62] and *SOX4* was reported to be necessary for tumor development driven by *PTEN* loss in a prostate cancer mouse model [118]. Likewise, several studies have shown that RNAi-mediated silencing of *SOX4 in vitro* or in an *in vivo* mouse model results in G0/G1 cell cycle arrest and leads to decreased cell proliferation, increased apoptosis and altered cell migration [119, 120].

SOX4 is known to regulate several key oncogenic signaling pathways in breast cancer including TGFβ, Wnt/β-catenin and PI3K. The effect of *SOX4* on TGFβ largely contributes to its role in regulating Epithelial-to-Mesenchymal Transition (EMT), cell migration, and metastasis; these aspects of *SOX4* activity will be discussed in greater detail in subsequent sections. However, a number of studies have shown that SOX4 and TGFβ can create a regulatory loop where SOX4 regulates and can be regulated by TGFβ activity. *SOX4* overexpression in MCF10A cells was shown to increased *TGFβ1* and *TGFβ2* expression leading to activation of TGFβ signaling as evident by increased phosphorylated SMAD2 levels; silencing of *SOX4* had the opposite effect on down-stream TGFβ signaling [62]. Likewise, SOX4 was found to bind to and activate down-stream components of the TGFβ pathway, including SMAD2 and SMAD3 in human mammary epithelial cells. In fact, SOX4 was found to co-localize with SMAD3 at multiple sites involved in metastasis, suggesting that these interactions may contribute to SOX4 mediated migration [121]. Interestingly, TGFβ was also shown to stimulate *SOX4* expression in murine breast cancer cells. In this study, Tiwari and colleagues suggest that the observed increase in *SOX4* expression in response to TGFβ activity may occur through non-conical (SMAD-independent) signaling, possibly through Wnt and Notch activity [120]. These data collectively suggest that the interplay between SOX4 and TGFβ signaling is more complicated than initially believed and considerably more work will be required to fully delineate this signaling network in breast cancer.

Beyond TGFβ signaling, SOX4 has been implicated in regulating a number of other oncogenic signaling networks in breast cancer including the PI3 kinase [54] and Wnt/β-catenin [122] pathways. Previous studies, including those from the TCGA project and our own work, have reported increased and uniform activation of PI3K signaling in basal-like tumors [3, 54, 123, 124]. This pathway mediates multiple oncogenic processes including proliferation, metabolism, motility and genome instability [125–128]. Our laboratory performed an integrated proteogenomic analysis of more than 3,000 human breast tumors and identified increased DNA amplification frequency and mRNA expression of *SOX4* in human breast tumors that had high levels of PI3K activity [54]. Analysis of proteomic data from a subset of more than 700 breast tumors further confirmed that *SOX4* DNA copy number status correlated with protein and phosphoprotein expression of down-stream components of the PI3K/Akt pathway [54]. Importantly, of those tumors that showed *SOX4* amplification and overexpression, the vast majority were found to be TNBC or basal-like tumors. Finally, we validated these *in silico* findings through *in vitro* studies that demonstrated that siRNA-mediated silencing of *SOX4* resulted in a reduced Akt phosphorylation in TNBC or basal-like cell lines with high *SOX4* expression and high PI3K activity [54]. While the exact mechanisms by which SOX4 mediates PI3K signaling in TNBC remain to be elucidated, these findings are consistent with previous studies demonstrating that SOX4 can mediate Akt activity in prostate cancer and lymphoma models through tissue-specific mechanisms [118, 129]. The Wnt signaling pathway has also been found to play an important role in breast cancer development [130]. Previous studies have shown that β-catenin nuclear localization is significantly enriched in TNBC cell lines and tumors, indicating activation of Wnt signaling in these cells [119, 131]. Interestingly, SOX4 has been found to stabilize and prevent β-catenin from proteosomal degradation by upregulating expression of casein kinase 2 (*CK2*) in colon adenocarcinoma cells, suggestive of a possible mechanism by which SOX4 induces Wnt/β-catenin signaling in the context of breast cancer [131].

Consistent with the noted overexpression of *SOX2* in breast cancers, Chen *et al*. reported that SOX2 promotes cell proliferation and tumorigenesis. Evidence indicates that SOX2 mediates this effect in breast cancer by accelerating the G1/S transition of the cell cycle, in part, through activation of the Wnt/β-catenin pathway [132]. This signaling pathway is critical for several aspects of tumorigenesis and functions by stabilizing and accumulating β-catenin in the nucleus where it interacts with TCF/LEF transcription factors to activate down-stream target genes [130]. It was recently reported that β-

catenin interacts with SOX2 and this interaction mediates SOX2 DNA binding and transcriptional activity in breast cancer cells. Notably, it was observed that SOX2 is required to interact with β-catenin to mediate Cyclin D1 (*CCND1*) expression in order to modulate accelerated G1/S transition [132]. However, additional studies have provided confounding data regarding the impact of the interaction between SOX2 and β-catenin. Recent studies by Ye *et al*. demonstrated that the nuclear interaction between SOX2 and β-catenin only occurred in a subset of breast cancer cells that are responsive to SOX2 activity. This study showed that in a small population of SOX2 responsive cells, β-catenin interacted with and suppressed SOX2 activity which led to decreased SOX2-dependent mammosphere formation. RNAi mediated knockdown of β-catenin could rescue this effect indicating that β-catenin is an essential determinant of the DNA binding and transcriptional activity of *SOX2* [133]. Clearly additional studies will be required to fully delineate the differences between these studies and to elucidate the relationship between SOX2 and Wnt/β-catenin signaling in TNBC or basal-like breast cancers.

SOX2 has also been shown to control the expression of a number of microRNAs including miR-181a-5p and miR-30e-5p, both of which regulate SOX2-mediated oncogenesis by inhibiting the expression of Tumor Suppressor Candidate 3 (TUSC3) protein in breast cancer cells [134]. *TUSC3* expression has been shown to be negatively correlated with *SOX2* in human breast cancer samples and evidence indicates that upregulation of TUSC3 inhibits cell proliferation as well as the migration potential of breast cancer cells, suggesting that this may be a significant mechanism by which SOX2 mediates its effect on these processes [134]. While SOX2 can mediate cellular effects by regulating the expression of multiple miRNA, it was recently shown that miR-101 can inhibit *SOX2* activity and that overexpression of miR-101 resulted in inhibition of SOX2-mediated cell growth, proliferation, and migration and resulted in the induction of apoptosis in breast cancer cell lines [135].

In addition to increasing evidence delineating the mechanisms by which SOX4 and SOX2 mediate cellular signaling, proliferation and tumor growth, a limited but rapidly expanding literature has begun to report the impact of several other SOX proteins on these cellular processes in breast cancer. In many instances, investigators have found that many SOX family members alter similar cellular functions and, in some instance, utilize similar mechanisms to affect these processes.

As previously discussed, increased expression of the SOXE family members *SOX10* and *SOX9*, was identified in breast cancers, particularly basal-like and TNBCs, as well as in metastatic TNBCs and secretory carcinomas [23–25, 27, 30, 35, 37, 42–44]. *SOX10* has been reported to mediate proliferation through the Notch4-PBP-mediated pathway in mouse derived mammary epithelial cells in culture [136]. In basal-like breast cancer, SOX10 has been shown to upregulate expression of uridine diphosphate-galactose ceramide galactosyltransferase (*UGT8*), a key enzyme in sulfatide biosynthesis. This altered signaling resulted in activation of the integrin αVβ5 signaling which has been shown to promote tumor progression [137]. Likewise, *SOX9* has also been implicated as an oncogene and a key regulator of stemness in TNBC [33]. Similar to other SOX proteins, *SOX9* expression can also be regulated by multiple miRNA including miR-133b [138] and miR-511 [139]. Of particular note, miR-133b was shown to modulate *SOX9* expression and regulate *SOX9*-mediated tumorigenesis including the metastatic phenotype [138]. Interestingly, Zhoa *et al*. demonstrated that miR-511 inhibits breast cancer cell proliferation by targeting the expression of *SOX9* and inactivating the PI3K/Akt signaling pathway [139]. Consistent with these data, *SOX9* inhibition mimicked the tumor suppressive function of miR-511 and reintroduction of *SOX9* resulted in activation of the PI3K signaling network [139]. These data suggest that *SOX9*, similar to *SOX4*, may play an integral role in regulating activation of PI3K signaling observed in basal-like or TNBC tumors [138, 139].

While few studies have investigated the role of *SOX5* in breast cancer, evidence in other tissue types such as lung adenocarcinoma [140] and osteosarcoma [141] suggests that this gene may be essential for breast cancer growth and progression, including metastasis. *SOX5* was shown to be significantly upregulated in TNBC cell lines and *in vitro* studies clearly demonstrated that this protein regulates breast cancer cell proliferation [39]. Similar to other SOX proteins, *SOX5* expression can be regulated by miRNA (miR-146a-5p) in TNBC clinical specimens and cell lines [142]. This appears to be a significant clinical association and is likely essential for regulating *SOX5* activity since miR-146a-5p is expressed at low levels in TNBC tumors [142]. Given that, *in vitro* studies demonstrated that overexpression of miR-146a-5p in breast cancer cell lines inhibit *SOX5*-induced cell proliferation [142], these data suggest that the interplay between miR-146a-5p and *SOX5* may be essential for regulating TNBC growth in a subset of these tumors. These findings are consistent with the previous studies demonstrating that SOX5 can regulate cell proliferation in lung adenocarcinoma and osteosarcoma by mediating the G1/S cell cycle transition [140, 141].

Finally, a number of studies have demonstrated that *SOX11*, *SOX12*, and *SOX18* are overexpressed in human breast cancers [32, 34, 90, 143, 144]. In each instance, experimental evidence indicates that these proteins mediate proliferation, migration, invasion and induction of apoptosis in both *in vitro* and *in vivo* models of breast cancer. However, a

limited number of studies have investigated oncogenic signaling mechanisms regulated by these factors in breast cancer suggesting that this will be an area of interest for future studies given the impact of these factors on transformation and tumorigenesis in other tissue types.

In contrast to the oncogenic properties demonstrated by the majority of the SOX proteins, a number of studies have determined that members of the SOXF family (*SOX7* and *SOX17*) as well as *SOX1* function as tumor suppressors in breast cancer [31, 38, 40, 46, 70, 73]. Interestingly, little has been reported about the mechanisms by which these proteins mediate their cellular functions in breast cancer. As previously noted, these proteins are significantly down-regulated in breast cancer cell lines and tissue samples [31, 38, 40, 46, 71, 73]. Consistent with these findings, ectopic overexpression of *SOX1* has been shown to prevent cell proliferation and invasion and induce apoptosis in breast cancer cells [31]. Mechanistically, *SOX1* overexpression results in repression of *CTNNB1* (β-catenin), *CCND1* (Cyclin D1) and *MYC* expression, suggesting that the tumor suppressive properties of SOX1 are mediated in part by regulating the Wnt/β-catenin pathway [31].

Both *SOX7* and *SOX17* of the SOXF family have been shown to be tumor suppressor proteins. Similar to *SOX1*, both proteins have been found to inhibit activity of Wnt/β-catenin signaling. In breast cancer cell lines and tumor samples, *SOX17* expression is epigenetically inactivated by promoter methylation and was found to be negatively correlated with Wnt/β-catenin signaling [46]. *SOX17* methylation status was associated with higher tumor grade, lymph node metastasis and shorter disease-free and overall survival compared to normal *SOX17* expression. Treatment with 5-aza-2'-deoxycytidine, a demethylating agent, restored *SOX17* expression at both protein and RNA levels. This was associated with a significant reduction in β-catenin levels in breast cancer cell lines suggesting that *SOX17* promoter hypermethylation leads to aberrant activation of Wnt/β-catenin signaling in breast cancer cells [38, 46].

In addition to promoter methylation, *SOX17* expression is also subject to regulation by microRNAs. Yang *et al*. showed that miR-194-5p promoted cell proliferation, migration and invasion in TNBC cell lines by suppressing *SOX17* expression and activating the Wnt/β-catenin signaling pathway [71]. While miR-194-5p acts an oncogene, miR-340 was shown to act as a tumor suppressor by positively regulating the expression of *SOX17* and retinoblastoma (Rb) protein and negatively regulating the expression of *SOX2* in TNBC cell lines [68].

Like *SOX17*, ectopic expression of *SOX7* decreased cellular proliferation, metastasis and *in vivo* growth, while inhibiting the expression of this gene enhanced these cellular functions [40, 70]. Similar to *SOX17*, *SOX7* can be regulated by over-expression of an oncogenic microRNA. Shen *et al*. demonstrated that ectopic expression of miR-492 led to increased proliferation and the upregulation of *CCND1* (Cyclin D1) and *MYC* through suppression of *SOX7* activity [69].

Collectively these studies suggest that oncogenic and tumor suppressive SOX proteins activate and repress many of the same signaling pathways in TNBC or basal-like breast cancer. As such, understanding the interplay between these proteins in tumor development and progression as well as the mechanisms by which these proteins regulate various signaling networks will be necessary to clarify SOX-mediated tumorigenesis.

The Role of SOX Proteins on EMT, Migration and Metastasis

Epithelial-to-mesenchymal transition (EMT) is a highly complex and orchestrated trans-differentiation process that takes place during development and tumorigenesis and involves the depolarization of epithelial cells into a highly invasive mesenchymal phenotype [145]. EMT is commonly associated with progression of malignancy and tumor metastasis, characterized by invasion and migration of primary tumor cells into distant sites [145]. Consistent with their role in development, SOX family members have been shown to play an integral role in regulating EMT as well as tumor migration and invasion.

SOX2 expression has been shown to be significantly upregulated in early stage and metastatic breast carcinoma [45, 47, 66, 67]. Consistent with these studies, a causal link between high *SOX2* expression and EMT has also been established in breast cancer. Overexpression of *SOX2* in breast cancer cells lines was shown to induce EMT by activating the Wnt/β-catenin signaling pathway [146]. Pang *et al.* demonstrated that miR-200 targets *SOX2* in TNBC cells and inhibits migration, invasion and mammosphere formation in these cells [147]. These investigators further demonstrated that MYC recruits DNMT3A to the miR-200 promoter in order to promote CpG island hypermethylation and subsequent repression of miR-200 expression [147]. Interestingly, miR-101 was also shown to inhibit EMT in breast cancer cells by directly targeting *SOX2* expression [135].

SOX4 has also been reported to play a critical role in the regulation of EMT in breast cancer [62, 120, 148]. Constitutive *SOX4* expression in a mammary epithelial cell line resulted in the induction of a mesenchymal phenotype associated with a decrease in E-cadherin and β-catenin and increase in N-cadherin and vimentin protein and mRNA expression [62]. Interestingly, Tiwari *et al*. demonstrated that SOX4 regulates EMT through epigenetic reprogramming by targeting Ezh2, part of the Polycomb repressor complex 2. The authors demonstrated that Ezh2 reprograms the

epigenome by promoting H3K27me3 repressive marks on the promoter region of epithelial genes and silences them to induce EMT [120]. Consistent with these findings, increased *SOX4* expression has been shown to be associated with invasive cancer subtypes and higher tumor grades [62, 120]. Furthermore, the increase in mesenchymal markers following *SOX4* overexpression was dependent on activation of the TGFβ pathway mediated by SOX4 [62, 120, 148]. Activation of TGFβ signaling is a common event in human cancer progression and acts as major inducer of EMT [149]. As we have discussed previously, expression of *SOX4* also appears to be regulated in response to TGFβ, suggesting an auto-regulatory loop may dictate the expression and activation of these genes and signaling pathways in breast cancer [62, 120]. While the association between SOX4 and TGFβ appears to be well established, additional studies are clearly required to fully delineate the mechanisms by which each signaling pathway regulates the other and the impact of these relationships in breast cancer.

Consistent with its role in regulating EMT, *SOX4* has been shown to enhance tumor invasion in multiple tissue types including breast, ovarian, prostate, melanoma, hepatocellular carcinoma and lung cancer through tissue-specific mechanisms [57, 118, 150–154]. In breast cancer, Tavazoie *et al.* demonstrated that shRNA-mediated inhibition of *SOX4* in a highly metastatic derivative of MDA-MB-231 (MDA-LM2) cells resulted in decreased lung metastasis in xenograft mouse models [151]. More recently, Lee *et al.* showed that SOX4 activated *TMEM2* gene expression in MDA-LM2 cells which promoted tumor cell migration and invasion [155]. In agreement with these findings, Tiwari and colleagues demonstrated that depletion of *SOX4* in a murine breast cancer cell line derived tumor model resulted in loss of tumor metastasis. In this study, investigators showed that ablation of *SOX4* in Py2T cell lines resulted in decreased primary tumor growth and metastatic spread to the axillary and inguinal lymph nodes, lungs and livers of nude mice following subcutaneous transplantation [120]. These studies clearly indicate the essentiality of *SOX4* in regulating EMT and promoting tumor invasion in breast cancer.

SOX4 mediated induction of EMT in breast cancer cells is also subjected to regulation by microRNAs. *SOX4* is a direct target of miR-93, which downregulates proliferation and differentiation in breast cancer cells, as well as normal breast stem cells isolated from reduction surgeries [156]. Ectopic expression of miR-93 in breast cancer cells reversed the process of EMT by inducing the mesenchymal-to-epithelial transition (MET) associated with the loss of TGFβ signaling and decrease in cancer stem cell population as well as a reduction in *in vivo* tumor development and metastasis [156]. miR-212/miR-132, miR-338-3p and miR-320, direct targets of SOX4, have been shown to mediate tumor suppressive effects as overexpression of these miRs resulted in decreased migration and invasion while inhibiting their expression induced these SOX4-driven phenotypes [157–159].

Beyond direct regulation of *SOX4*-mediated EMT and cell migration by miRNAs, additional studies have indicated that *SOX4* can be regulated by androgen receptor and CXCL1 activity [160, 161]. These studies indicate that androgen receptor activity repressed the expression of the long non-coding RNA, *ARNILA* which has been shown to correlate with EMT, poor progression-free survival, and metastasis [161]. Importantly, it was reported that *ARNILA* could bind to miR-204 which resulted in suppression of miR-204 activity and led to rescue of *SOX4* expression in TNBC cells lines. Furthermore, *ARNILA*-mediated rescue of *SOX4* expression promoted invasion and metastasis *in vitro* as well as *in vivo* [161]. Likewise, CXCL1, a cytokine secreted by tumor associated macrophages, which is highly expressed in breast cancer lung metastases, was shown to upregulate *SOX4* mRNA and protein expression levels in tumor epithelial cells [160]. In this study, investigators demonstrated that CXCL1 treatment led to increased enrichment of NFκB at the *SOX4* promoter resulting in increased *SOX4* expression. Consistent with these findings, inhibiting the NFκB pathway using BAY-11-7082 was found to repress SOX4 activity and inhibit the EMT process modulated by CXCL1 [160].

As previously discussed, *SOX5* expression has been reported to be significantly upregulated in TNBC cell lines and shown to mediate breast cancer cell proliferation, migration and invasion through induction of EMT [39]. Recent studies have indicated that SOX5 is significantly enriched on the *TWIST1* promoter, thereby directly regulating its expression. Importantly, lentiviral mediated silencing of *SOX5* expression inhibited SOX5-mediated EMT, thereby suppressing the oncogenic activity of the SOX5 protein [39]. Interestingly, a number of subsequent studies have indicated that SOX5 can similarly mediate EMT and cell migration through TGFβ activity and/or Twist or Snail expression in lung, prostate, pituitary and hepatocellular carcinoma [140, 162–164]. Additional studies suggest that changes in specific miRNA, including miR-139-5p, miR-132, miR-15a, and miR-16 may regulate *SOX5* expression and SOX5-mediated EMT [163, 165]. These data suggest that SOX5 and SOX5-mediated EMT may be similarly regulated in breast cancer or TNBC; however additional studies will be essential to delineate these regulatory mechanisms.

Finally, a number of studies have implicated additional oncogenic SOX proteins including *SOX9* [138, 139], *SOX10* [88, 89], *SOX11* [32] and *SOX12* [34] in the regulation of EMT, cell migration and metastasis; however additional insight into the mechanism(s) by which these proteins regulate these processes in breast cancer is needed. SOX9 has been reported to regulate EMT in TNBC cell lines and tissue samples [138, 139]. While the exact mechanisms by which SOX9 mediates these processes remains unclear, *SOX9* activity has

been shown to be regulated by miR-206 as well as miR-113b and miR-511 [138, 139, 166]. Multiple studies have indicated that loss of these miRNAs leads to increased *SOX9* expression and results in increased invasion and migration. Limited studies do suggest that SOX9 may partially mediate this effect on cancer cells via the PI3K/Akt signaling network [138, 139]; however additional studies will be required to fully delineate these mechanisms. Similarly, *SOX10* was found to be highly expressed in fetal mammary stem cells (fMaSCs), which closely resemble the MaSC-like cancer cells in breast tumors, and *SOX10* overexpression in primary organoids resulted in activation of EMT and migration of cells away from the primary organoids, suggesting that SOX10 may be important for tumor cell spread [75, 88, 89]. Consistent with this idea, ectopic expression of *SOX10* was shown to induce vimentin, Snail2, and Twist1 expression which resulted in increased EMT in breast cancer models [88]. Lastly, SOX12 was also found to regulate the migratory and invasive phenotypes in breast cancer cell line studies, although the mechanism remains unclear [34].

In summary, the evidence clearly indicates the important role of SOX family members in promoting breast cancer cell motility, including changes in cell morphology, migration and invasion. SOX4, SOX2 and SOX5 have been the best studied of these proteins as evidence is available of the mechanisms by which they mediate phenotypic changes to enhance breast cancer metastasis and TGFβ and Wnt/β-catenin signaling pathways seems to be the predominant signaling pathways through which SOX proteins mediate EMT. Further studies are needed to better understand these mechanisms as well as elucidate the role of other SOX family members in promoting EMT and metastasis.

Therapeutic Potential of Sox Signaling in Breast Cancer

Given the noted oncogenic role of SOX family proteins in development and breast cancer tumorigenesis, these genes, or the down-stream signaling pathways activated by these genes, represent potential therapeutic targets for breast cancer treatment. In fact, a number of investigators have begun to address these questions. Compounds that are specific for SOX2 and SOX18 inhibition have been identified and shown to have potential as therapeutic agents.

A high-throughput fluorescence anisotropy screen identified a Dawson polyoxometalates (POMs) as a specific inhibitor of SOX2. This compound, $K_6P_2Mo_{18}O_{62}$, was found to be highly selective to the DNA binding domain of SOX2 and related HMG-containing proteins at a nanomolar concentration. Importantly, experimental evidence demonstrated that this compound inhibited the ability of SOX2 to bind DNA. While additional studies would be required to alleviate concerns regarding selectivity, the potential clearly exists for this compound or its subsequent derivatives to be utilized as a framework for the development of future anticancer therapeutics targeting the similar DNA binding domain of transcription factors of the SOX protein family [167].

Furthermore, Overman *et al.* used a combination of genomic, proteomic and biophysical methods to identify a panel of protein–protein interactions that may be essential for regulating SOX18 activity. The investigators utilized a natural small molecule inhibitor, Sm4, to specifically target these interactions [26]. Pharmacological inhibition of SOX18 using Sm4 significantly increased the overall survival of BALB/c mice that expressed tumors derived from aggressive and highly metastatic 4T1.2 mammary carcinoma cells [26]. Although this compound had no effect on the size of the primary tumor, a significant reduction in the number of lung metastases was observed. This effect was attributed to the reduction in tumor induced angiogenesis as demonstrated by an overall reduction in the volume of blood vessels in the tumors of Sm4 treated animals [26]. Consistent with these findings, it was recently reported that SOX4 mediates angiogenesis in breast cancer by directly regulating expression of endothelin-1 (ET-1) [61]. Given that additional SOX proteins have been shown to play a role in regulating angiogenesis in lung adenocarcinoma [168] and melanoma [169], these results suggest that SOX-

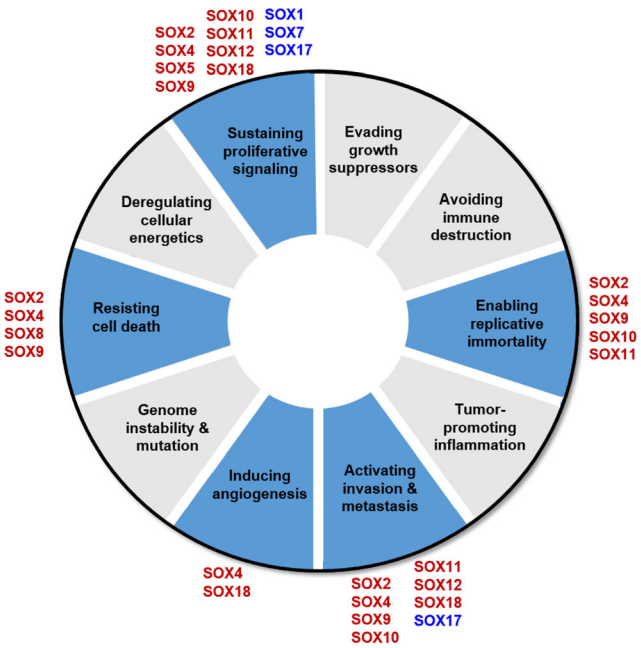

Fig. 3 Schematic overview of phenotypic functions regulated by SOX proteins in breast cancer. The hallmarks of cancer regulated by SOX proteins in breast cancer are represented. The hallmarks that are specifically shown to be regulated by SOX proteins are highlighted in blue while those that have not been reported to be affected by SOX proteins are indicated in gray. Individual SOX proteins that have been reported to activate (red) or repress (blue) each of these hallmarks in the context of breast cancer are indicated

mediated angiogenesis, either directly or indirectly, may represent a potential therapeutic opportunity. Future research into understanding the interacting partners of SOX proteins as well as the mechanisms of SOX-mediated angiogenesis has the potential to aid in the identification of novel therapeutic targets to inhibit SOX protein activity and/or SOX-mediated angiogenesis.

Dynamic epigenetic regulation by DNA methylation and histone modification by chromatin remodeling proteins results in an altered and reversible epigenetic landscape. Emerging evidence indicates that cancer cells can become addicted to the aberrantly developed epigenetic landscape in a manner similar to the dependency of tumor cells on specific oncogenes (*i.e.* oncogene addiction). These data suggest that cancer cells that are dependent on this altered epigenomic landscape would be more sensitive than normal cells to epigenetic therapy. However, inhibitors targeting DNA methyltransferase (DNMT) enzymes involved in the silencing of tumor suppressive SOX proteins, SOX7 and SOX17, have shown limited promise in hematological malignancies due to their lack of specificity as well as cytotoxicity resulting from global hypermethylation [170, 171]. Moreover, 3-deazaneplanocin A (DZNep), a drug targeting EZH2 which is directly regulated by SOX4 in breast cancer, shows limited specificity towards EZH2 and inactivates multiple histone methyltransferases resulting in aberrant reactivation of the developmental genes in cancer cells [172, 173]. Thus, although epigenetic therapies hold great promise for development of anticancer therapies, additional studies are required to address concerns regarding specificity and cytotoxicity that limit the development of rationally defined epigenetic therapies for breast tumors.

Given these results, there is a clear potential to develop novels strategies to regulate these genes directly or indirectly through dependent co-factors or down-stream targets including TGFβ and PI3K-family targeting drugs. While many of the SOX proteins have been shown to be essential for cell survival, tumor cell proliferation and growth, as well as a multitude of other tumor characteristics, a number of limitations must be recognized in considering these genes as potential drug targets. To begin, with few exceptions, transcription factors are notoriously difficult to inhibit therapeutically. While alternative strategies may exist, as we have outlined above, developing therapeutic approaches incorporating the inhibition of SOX proteins may prove to be difficult beyond technical challenges. Most notably, as we have outlined in this review, a number of functional redundancies exist between different SOX family members which may limit the ability to develop compounds that directly target any given SOX

Table 1 Summary of SOX protein function and expression in breast cancer

Group	Gene	Chromosomal position	Role in breast cancer
SOXB	SOX 1	13q34	Tumor suppressor and decreased expression in breast cancer, suppresses oncogenic Wnt/ β-catenin pathway by downregulating β-catenin, cyclin Dl and Myc [29].
	SOX2	3q26.33	Embryonic and stem cell factor [81–83], highly expressed in BLBC and breast cancer stem cells [43, 45], regulates oncogenic phenotypes [123, 125].
SOXC	SOX4	6p22.3	Essential embryonic transcription factor with high expression in mammary glands [48, 49], stem cell factor and positively regulates EMT [58, 110, 111, 150], amplified and increased expression in BLBC and TNBC, regulates PI3K/Akt, TGF-β and WNT signaling, as well as positively correlates with metastasis and tumor volume [50, 51, 57].
	SOX11	2p25.2	Essential factor for regulating oncogenic and stem cells phenotypes in BLBC [30, 80]
	SOX12	20p13	Increased expression in breast cancer, regulates EMT and proliferation in vitro and in vivo [32].
SOXD	SOX5	12p12.1	Increased expression in breast cancer, regulates proliferation, metastasis and invasion by regulating Twist1 expression [37].
SOXE	SOX8	16p13.3	Amplified in TNBC and negatively correlated with overall patient survival [60].
	SOX9	17q23	Embryonic and stem cell factor [62, 97, 98], overexpression converts luminal cells into MaSCs [77], overexpressed in TNBC and associated with poor prognosis [31, 34].
	SOX10	22q13.1	Increased expression in basal-like and TNBCs [21, 22, 28, 41, 42], regulates oncogenic phenotypes by upregulating Vimentin, Snail2 and Twist1 expression [78, 79, 127, 128], regulates cancer stem cell phenotypes [33, 78, 79, 105, 106].
SOXF	SOX7	8p23.1	Tumor suppressor and decreased expression in tumors due to promoter hypermethylation [38, 136].
	SOX17	8q11.23	Tumor suppressor with decreased expression in breast cancer due to promoter hypermethylation, suppresses oncogenic Wnt/ βcatenin pathway and negatively correlates with tumor grade, lymph node metastasis, disease-free and overall survival [36, 44].
	SOX18	20q13.33	Increased expression in breast cancer and positively correlates with higher tumor grade. Regulates angiogenesis and pharmacological inhibition of SOX18 decreases lung metastases by reducing the volume of blood vessels in tumors [24, 39].

protein or SOX family of proteins. However, given the role that these proteins play in regulating tumor development and progression, additional studies have the potential to uncover novel approaches to directly or indirectly inhibit the impact of these proteins in breast cancer.

Concluding Remarks

It is clear that substantial progress has been made in the past few years in illuminating the role SOX proteins play in regulating important aspects of breast cancer genesis, stemness, development and therapeutic resistance (Fig. 3 and Table 1). Although SOX proteins are known to regulate other important hallmarks of cancer including evasion of growth suppressors [174, 175], immune modulation [176], deregulation of cellular energetics [177] and inflammatory processes [178] in other tissue types, additional studies are needed to assess the role of SOX proteins in regulation of these phenotypic hallmarks in breast cancer (Fig. 3). It is also apparent from the literature that the activating or repressing functions of SOX proteins in the developmental processes is highly dependent on their interacting protein partners, either transcription factors or epigenetic machinery and thus demonstrate high levels of tissue specificity [extensively reviewed in [18, 20]]. However, with respect to its function in breast cancer, it is still unclear if any of the SOX protein functional domains, lying outside the HMG DNA binding domains have any crucial roles in regulating key aspects of mammary tumorigenesis. Future research will provide more insight into the interactome and gene regulatory networks of SOX proteins that operate in the context of breast cancer. These findings will no doubt aid in the development of novel treatment strategies for this highly heterogeneous disease with limited therapeutic options.

Acknowledgements While we have attempted to include all relevant studies in this manuscript, we apologize to those investigators whose work was not included due to spatial limitations. We would like to thank the members of our lab for critical reading of this manuscript. This work is funded by post-doctoral fellowship DHFS-17PCC-002 from the New Jersey Commission for Cancer Research to G.A.M. and from the National Cancer Institute of the National Institutes of Health (CA166228), the V Foundation for Cancer Research (V2016-013), and the New Jersey Health Foundation (PC-17-18) to M.L.G.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018;68(1):7–30.
2. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869–74.
3. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70.
4. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346–52.
5. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, et al. Comprehensive molecular portraits of invasive lobular breast cancer. Cell. 2015;163(2):506–19.
6. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747–52.
7. Gatza ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, et al. A pathway-based classification of human breast cancer. Proc Natl Acad Sci U S A. 2010;107(15):6994–9.
8. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. 2010;12(5):R68.
9. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest. 2011;121(7):2750–67.
10. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature. 2016;534(7605):55–62.
11. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A. 2003;100(14):8418–23.
12. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, et al. Concordance among gene-expression-based predictors for breast cancer. N Engl J Med. 2006;355(6):560–9.
13. Gatza ML, Silva GO, Parker JS, Fan C, Perou CM. An integrated genomics approach identifies drivers of proliferation in luminal-subtype human breast cancer. Nat Genet. 2014;46(10):1051–9.
14. Prat A, Fan C, Fernandez A, Hoadley KA, Martinello R, Vidal M, et al. Response and survival of breast cancer intrinsic subtypes following multi-agent neoadjuvant chemotherapy. BMC Med. 2015;13:303.
15. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell. 2014;158(4):929–44.
16. Gross K, Wronski A, Skibinski A, Phillips S, Kuperwasser C. Cell fate decisions during breast cancer development. J Dev Biol. 2016;4(1):4.
17. Zhang M, Lee AV, Rosen JM. The cellular origin and evolution of breast cancer. Cold Spring Harb Perspect Med. 2017 Mar 1;7(3):a027128.
18. Kamachi Y, Kondoh H. Sox proteins: regulators of cell fate specification and differentiation. Development. 2013;140(20):4129–44.
19. Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. Cell Stem Cell. 2013;12(1):15–30.
20. She ZY, Yang WX. SOX family transcription factors involved in diverse cellular events during development. Eur J Cell Biol. 2015;94(12):547–63.
21. Gubbay J, Collignon J, Koopman P, Capel B, Economou A, Munsterberg A, et al. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature. 1990;346(6281):245–50.

22. Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, et al. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature. 1990;346(6281):240–4.
23. Tozbikian GH, Zynger DL. A combination of GATA3 and SOX10 is useful for the diagnosis of metastatic triple negative breast cancer. Hum Pathol. 2019;85:221-227.
24. Al-Zahrani KN, Cook DP, Vanderhyden BC, Sabourin LA. Assessing the efficacy of androgen receptor and Sox10 as independent markers of the triple-negative breast cancer subtype by transcriptome profiling. Oncotarget. 2018;9(70):33348–59.
25. Zang H, Li N, Pan Y, Hao J. Identification of upstream transcription factors (TFs) for expression signature genes in breast cancer. Gynecol Endocrinol. 2017;33(3):193–8.
26. Overman J, Fontaine F, Moustaqil M, Mittal D, Sierecki E, Sacilotto N, et al. Pharmacological targeting of the transcription factor SOX18 delays breast cancer in mice. Elife. 2017;6:e21221.
27. Nelson ER, Sharma R, Argani P, Cimino-Mathews A. Utility of Sox10 labeling in metastatic breast carcinomas. Hum Pathol. 2017;67:205–10.
28. Min L, Zhang C, Qu L, Huang J, Jiang L, Liu J, et al. Gene regulatory pattern analysis reveals essential role of core transcriptional factors' activation in triple-negative breast cancer. Oncotarget. 2017;8(13):21938–53.
29. Feng X, Lu M. Expression of sex-determining region Y-box protein 2 in breast cancer and its clinical significance. Saudi Med J. 2017;38(7):685–90.
30. Feng W, Liu S, Zhu R, Li B, Zhu Z, Yang J, et al. SOX10 induced Nestin expression regulates cancer stem cell properties of TNBC cells. Biochem Biophys Res Commun. 2017;485(2):522–8.
31. Song L, Liu D, He J, Wang X, Dai Z, Zhao Y, et al. SOX1 inhibits breast cancer cell growth and invasion through suppressing the Wnt/beta-catenin signaling pathway. APMIS. 2016;124(7):547–55.
32. Shepherd JH, Uray IP, Mazumdar A, Tsimelzon A, Savage M, Hilsenbeck SG, et al. The SOX11 transcription factor is a critical regulator of basal-like breast cancer growth, invasion, and basal-like gene expression. Oncotarget. 2016;7(11):13106–21.
33. Lei B, Zhang YX, Liu T, Li YW, Pang D. Sox9 upregulation in breast cancer is correlated with poor prognosis and the CD44(+)/CD24(-/low) phenotype. Int J Clin Exp Pathol. 2016;9(7):7345–51.
34. Ding H, Quan H, Yan W, Han J. Silencing of SOX12 by shRNA suppresses migration, invasion and proliferation of breast cancer cells. Biosci Rep. 2016;36(5):e00389.
35. Ye X, Tam WL, Shibue T, Kaygusuz Y, Reinhardt F, Ng Eaton E, et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. Nature. 2015;525(7568):256–60.
36. Pomp V, Leo C, Mauracher A, Korol D, Guo W, Varga Z. Differential expression of epithelial-mesenchymal transition and stem cell markers in intrinsic subtypes of breast cancer. Breast Cancer Res Treat. 2015;154(1):45–55.
37. Miettinen M, McCue PA, Sarlomo-Rikala M, Biernat W, Czapiewski P, Kopczynski J, et al. Sox10–a marker for not only schwannian and melanocytic neoplasms but also myoepithelial cell tumors of soft tissue: a systematic analysis of 5134 tumors. Am J Surg Pathol. 2015;39(6):826–35.
38. Fu D, Ren C, Tan H, Wei J, Zhu Y, He C, et al. Sox17 promoter methylation in plasma DNA is associated with poor survival and can be used as a prognostic factor in breast cancer. Medicine (Baltimore). 2015;94(11):e637.
39. Pei XH, Lv XQ, Li HX. Sox5 induces epithelial to mesenchymal transition by transactivation of Twist1. Biochem Biophys Res Commun. 2014;446(1):322–7.
40. Stovall DB, Wan M, Miller LD, Cao P, Maglic D, Zhang Q, et al. The regulation of SOX7 and its tumor suppressive role in breast cancer. Am J Pathol. 2013;183(5):1645–53.
41. Pula B, Olbromski M, Wojnar A, Gomulkiewicz A, Witkiewicz W, Ugorski M, et al. Impact of SOX18 expression in cancer cells and vessels on the outcome of invasive ductal breast carcinoma. Cell Oncol (Dordr). 2013;36(6):469–83.
42. Mohamed A, Gonzalez RS, Lawson D, Wang J, Cohen C. SOX10 expression in malignant melanoma, carcinoma, and normal tissues. Appl Immunohistochem Mol Morphol. 2013;21(6):506–10.
43. Ivanov SV, Panaccione A, Nonaka D, Prasad ML, Boyd KL, Brown B, et al. Diagnostic SOX10 gene signatures in salivary adenoid cystic and breast basal-like carcinomas. Br J Cancer. 2013;109(2):444–51.
44. Cimino-Mathews A, Subhawong AP, Elwood H, Warzecha HN, Sharma R, Park BH, et al. Neural crest transcription factor Sox10 is preferentially expressed in triple-negative and metaplastic breast carcinomas. Hum Pathol. 2013;44(6):959–65.
45. Leis O, Eguiara A, Lopez-Arribillaga E, Alberdi MJ, Hernandez-Garcia S, Elorriaga K, et al. Sox2 expression in breast tumours and activation in breast cancer stem cells. Oncogene. 2012;31(11):1354–65.
46. Fu DY, Wang ZM, Li C, Wang BL, Shen ZZ, Huang W, et al. Sox17, the canonical Wnt antagonist, is epigenetically inactivated by promoter methylation in human breast cancer. Breast Cancer Res Treat. 2010;119(3):601–12.
47. Rodriguez-Pinilla SM, Sarrio D, Moreno-Bueno G, Rodriguez-Gil Y, Martinez MA, Hernandez L, et al. Sox2: a possible driver of the basal-like phenotype in sporadic breast cancer. Mod Pathol. 2007;20(4):474–81.
48. Hunt SM, Clarke CL. Expression and hormonal regulation of the Sox4 gene in mouse female reproductive tissues. Biol Reprod. 1999;61(2):476–81.
49. Schilham MW, Oosterwegel MA, Moerer P, Ya J, de Boer PA, van de Wetering M, et al. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. Nature. 1996;380(6576):711–4.
50. Lefebvre V, Dumitriu B, Penzo-Mendez A, Han Y, Pallavi B. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int J Biochem Cell Biol. 2007;39(12):2195–214.
51. Prior HM, Walter MA. SOX genes: architects of development. Mol Med. 1996;2(4):405–12.
52. Chew LJ, Gallo V. The Yin and Yang of Sox proteins: activation and repression in development and disease. J Neurosci Res. 2009;87(15):3277–87.
53. Thu KL, Becker-Santos DD, Radulovich N, Pikor LA, Lam WL, Tsao MS. SOX15 and other SOX family members are important mediators of tumorigenesis in multiple cancer types. Oncoscience. 2014;1(5):326–35.
54. Mehta GA, Parker JS, Silva GO, Hoadley KA, Perou CM, Gatza ML. Amplification of SOX4 promotes PI3K/Akt signaling in human breast cancer. Breast Cancer Res Treat. 2017;162(3):439–50.
55. Song GD, Sun Y, Shen H, Li W. SOX4 overexpression is a novel biomarker of malignant status and poor prognosis in breast cancer patients. Tumour Biol. 2015;36(6):4167–73.
56. Castillo SD, Matheu A, Mariani N, Carretero J, Lopez-Rios F, Lovell-Badge R, et al. Novel transcriptional targets of the SRY-HMG box transcription factor SOX4 link its expression to the development of small cell lung cancer. Cancer Res. 2012;72(1):176–86.
57. Liao YL, Sun YM, Chau GY, Chau YP, Lai TC, Wang JL, et al. Identification of SOX4 target genes using phylogenetic footprinting-based prediction from expression microarrays

suggests that overexpression of SOX4 potentiates metastasis in hepatocellular carcinoma. Oncogene. 2008;27(42):5578–89.
58. Liu P, Ramachandran S, Ali Seyed M, Scharer CD, Laycock N, Dalton WB, et al. Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. Cancer Res. 2006;66(8):4011–9.
59. Aaboe M, Birkenkamp-Demtroder K, Wiuf C, Sorensen FB, Thykjaer T, Sauter G, et al. SOX4 expression in bladder carcinoma: clinical aspects and in vitro functional characterization. Cancer Res. 2006;66(7):3434–42.
60. Lee CJ, Appleby VJ, Orme AT, Chan WI, Scotting PJ. Differential expression of SOX4 and SOX11 in medulloblastoma. J Neuro-Oncol. 2002;57(3):201–14.
61. Vervoort SJ, de Jong OG, Roukens MG, Frederiks CL, Vermeulen JF, Lourenco AR, et al. Global transcriptional analysis identifies a novel role for SOX4 in tumor-induced angiogenesis. Elife. 2018;7:e27706.
62. Zhang J, Liang Q, Lei Y, Yao M, Li L, Gao X, et al. SOX4 induces epithelial-mesenchymal transition and contributes to breast cancer progression. Cancer Res. 2012;72(17):4597–608.
63. Dong P, Yu B, Pan L, Tian X, Liu F. Identification of Key genes and pathways in triple-negative breast cancer by integrated bioinformatics analysis. Biomed Res Int. 2018;2018:2760918.
64. Kundig P, Giesen C, Jackson H, Bodenmiller B, Papassotirolopus B, Freiberger SN, et al. Limited utility of tissue micro-arrays in detecting intra-tumoral heterogeneity in stem cell characteristics and tumor progression markers in breast cancer. J Transl Med. 2018;16(1):118.
65. Chakravarty G, Moroz K, Makridakis NM, Lloyd SA, Galvez SE, Canavello PR, et al. Prognostic significance of cytoplasmic SOX9 in invasive ductal carcinoma and metastatic breast cancer. Exp Biol Med (Maywood). 2011;236(2):145–55.
66. Lengerke C, Fehm T, Kurth R, Neubauer H, Scheble V, Muller F, et al. Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma. BMC Cancer. 2011;11:42.
67. Liu P, Tang H, Song C, Wang J, Chen B, Huang X, et al. SOX2 promotes cell proliferation and metastasis in triple negative breast cancer. Front Pharmacol. 2018;9:942.
68. Mohammadi Yeganeh S, Vasei M, Tavakoli R, Kia V, Paryan M. The effect of miR-340 over-expression on cell-cycle-related genes in triple-negative breast cancer cells. Eur J Cancer Care (Engl). 2017;26(6):10.1111/ecc.12496.
69. Shen F, Cai WS, Feng Z, Li JL, Chen JW, Cao J, et al. MiR-492 contributes to cell proliferation and cell cycle of human breast cancer cells by suppressing SOX7 expression. Tumour Biol. 2015;36(3):1913–21.
70. Stovall DB, Cao P, Sui G. SOX7: from a developmental regulator to an emerging tumor suppressor. Histol Histopathol. 2014;29(4):439–45.
71. Yang F, Xiao Z, Zhang S. Knockdown of miR-194-5p inhibits cell proliferation, migration and invasion in breast cancer by regulating the Wnt/beta-catenin signaling pathway. Int J Mol Med. 2018;42(6):3355–63.
72. Liu H, Mastriani E, Yan ZQ, Yin SY, Zeng Z, Wang H, et al. SOX7 co-regulates Wnt/beta-catenin signaling with Axin-2: both expressed at low levels in breast cancer. Sci Rep. 2016;6:26136.
73. Katoh M. Expression of human SOX7 in normal tissues and tumors. Int J Mol Med. 2002;9(4):363–8.
74. Dalerba P, Cho RW, Clarke MF. Cancer stem cells: models and concepts. Annu Rev Med. 2007;58:267–84.
75. Spike BT, Engle DD, Lin JC, Cheung SK, La J, Wahl GM. A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. Cell Stem Cell. 2012;10(2):183–97.
76. Zvelebil M, Oliemuller E, Gao Q, Wansbury O, Mackay A, Kendrick H, et al. Embryonic mammary signature subsets are activated in Brca1-/- and basal-like breast cancers. Breast Cancer Res. 2013;15(2):R25.
77. Adorno-Cruz V, Kibria G, Liu X, Doherty M, Junk DJ, Guan D, et al. Cancer stem cells: targeting the roots of cancer, seeds of metastasis, and sources of therapy resistance. Cancer Res. 2015;75(6):924–9.
78. Aiello NM, Stanger BZ. Echoes of the embryo: using the developmental biology toolkit to study cancer. Dis Model Mech. 2016;9(2):105–14.
79. Zhao D, Pan C, Sun J, Gilbert C, Drews-Elger K, Azzam DJ, et al. VEGF drives cancer-initiating stem cells through VEGFR-2/Stat3 signaling to upregulate Myc and Sox2. Oncogene. 2015;34(24):3107–19.
80. Visvader JE, Lindeman GJ. Cancer stem cells: current status and evolving complexities. Cell Stem Cell. 2012;10(6):717–28.
81. Wahl GM, Spike BT. Cell state plasticity, stem cells, EMT, and the generation of intra-tumoral heterogeneity. NPJ Breast Cancer. 2017;3:14.
82. Piva M, Domenici G, Iriondo O, Rabano M, Simoes BM, Comaills V, et al. Sox2 promotes tamoxifen resistance in breast cancer cells. EMBO Mol Med. 2014;6(1):66–79.
83. Abdelalim EM, Emara MM, Kolatkar PR. The SOX transcription factors as key players in pluripotent stem cells. Stem Cells Dev. 2014;23(22):2687–99.
84. Grosschedl R, Giese K, Pagel J. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. Trends Genet. 1994;10(3):94–100.
85. Liu K, Lin B, Zhao M, Yang X, Chen M, Gao A, et al. The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. Cell Signal. 2013;25(5):1264–71.
86. Pece S, Tosoni D, Confalonieri S, Mazzarol G, Vecchi M, Ronzoni S, et al. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. Cell. 2010;140(1):62–73.
87. Guo W, Keckesova Z, Donaher JL, Shibue T, Tischler V, Reinhardt F, et al. Slug and Sox9 cooperatively determine the mammary stem cell state. Cell. 2012;148(5):1015–28.
88. Dravis C, Chung CY, Lytle NK, Herrera-Valdez J, Luna G, Trejo CL, et al. Epigenetic and transcriptomic profiling of mammary gland development and tumor models disclose regulators of cell state plasticity. Cancer Cell. 2018;34(3):466–82 e6.
89. Dravis C, Spike BT, Harrell JC, Johns C, Trejo CL, Southard-Smith EM, et al. Sox10 regulates stem/progenitor and mesenchymal cell states in mammary epithelial cells. Cell Rep. 2015;12(12):2035–48.
90. Oliemuller E, Kogata N, Bland P, Kriplani D, Daley F, Haider S, et al. SOX11 promotes invasive growth and ductal carcinoma in situ progression. J Pathol. 2017;243(2):193–207.
91. Stevanovic M, Zuffardi O, Collignon J, Lovell-Badge R, Goodfellow P. The cDNA sequence and chromosomal location of the human SOX2 gene. Mamm Genome. 1994;5(10):640–2.
92. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. 2003;17(1):126–40.
93. Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, et al. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol. 2007;9(6):625–35.
94. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663–76.
95. Dong C, Wilhelm D, Koopman P. Sox genes and cancer. Cytogenet Genome Res. 2004;105(2-4):442–7.
96. Weina K, Utikal J. SOX2 and cancer: current research and its implications in the clinic. Clin Transl Med. 2014;3:19.

97. Wuebben EL, Rizzino A. The dark side of SOX2: cancer - a comprehensive overview. Oncotarget. 2017;8(27):44917–43.
98. Gong X, Liu W, Wu L, Ma Z, Wang Y, Yu S, et al. Transcriptional repressor GATA binding 1-mediated repression of SRY-box 2 expression suppresses cancer stem cell functions and tumor initiation. J Biol Chem. 2018;293(48):18646–54.
99. Deng Z, Du WW, Fang L, Shan SW, Qian J, Lin J, et al. The intermediate filament vimentin mediates microRNA miR-378 function in cellular self-renewal by regulating the expression of the Sox2 transcription factor. J Biol Chem. 2013;288(1): 319–31.
100. Zhang Y, Eades G, Yao Y, Li Q, Zhou Q. Estrogen receptor alpha signaling regulates breast tumor-initiating cells by down-regulating miR-140 which targets the transcription factor SOX2. J Biol Chem. 2012;287(49):41514–22.
101. Picon-Ruiz M, Pan C, Drews-Elger K, Jang K, Besser AH, Zhao D, et al. Interactions between adipocytes and breast cancer cells stimulate cytokine production and drive Src/Sox2/miR-302b-mediated malignant progression. Cancer Res. 2016;76(2):491–504.
102. Chen L, Xiao Z, Meng Y, Zhao Y, Han J, Su G, et al. The enhancement of cancer stem cell properties of MCF-7 cells in 3D collagen scaffolds for modeling of cancer and anti-cancer drugs. Biomaterials. 2012;33(5):1437–44.
103. Feng S, Duan X, Lo PK, Liu S, Liu X, Chen H, et al. Expansion of breast cancer stem cells with fibrous scaffolds. Integr Biol (Camb). 2013;5(5):768–77.
104. Bhola NE, Balko JM, Dugger TC, Kuba MG, Sanchez V, Sanders M, et al. TGF-beta inhibition enhances chemotherapy action against triple-negative breast cancer. J Clin Invest. 2013;123(3): 1348–58.
105. Ikushima H, Todo T, Ino Y, Takahashi M, Saito N, Miyazawa K, et al. Glioma-initiating cells retain their tumorigenicity through integration of the Sox axis and Oct4 protein. J Biol Chem. 2011;286(48):41434–41.
106. Cheung M, Briscoe J. Neural crest development is regulated by the transcription factor Sox9. Development. 2003;130(23):5681–93.
107. Nowak JA, Polak L, Pasolli HA, Fuchs E. Hair follicle stem cells are specified and function in early skin morphogenesis. Cell Stem Cell. 2008;3(1):33–43.
108. Domenici G, Aurrekoetxea-Rodriguez I, Simoes BM, Rabano M, Lee SY, Millan JS, et al. A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. 2019; Oncogene 38:3151–3169.
109. Wang C, Christin JR, Oktay MH, Guo W. Lineage-biased stem cells maintain estrogen-receptor-positive and -negative mouse mammary luminal lineages. Cell Rep. 2017;18(12):2825–35.
110. Xue Y, Lai L, Lian W, Tu X, Zhou J, Dong P, et al. SOX9/FXYD3/Src axis is critical for ER(+) breast cancer stem cell function. Mol Cancer Res. 2019;17(1):238-249.
111. Jeselsohn R, Cornwell M, Pun M, Buchwalter G, Nguyen M, Bango C, et al. Embryonic transcription factor SOX9 drives breast cancer endocrine resistance. Proc Natl Acad Sci U S A. 2017;114(22):E4482–E91.
112. Mateo F, Arenas EJ, Aguilar H, Serra-Musach J, de Garibay GR, Boni J, et al. Stem cell-like transcriptional reprogramming mediates metastatic resistance to mTOR inhibition. Oncogene. 2017;36(19):2737–49.
113. Li Q, Yao Y, Eades G, Liu Z, Zhang Y, Zhou Q. Downregulation of miR-140 promotes cancer stem cell formation in basal-like early stage breast cancer. Oncogene. 2014;33(20):2589–600.
114. Giraddi RR, Chung CY, Heinz RE, Balcioglu O, Novotny M, Trejo CL, et al. Single-cell transcriptomes distinguish stem cell state changes and lineage specification programs in early mammary gland development. Cell Rep. 2018;24(6):1653–66 e7.
115. Panaccione A, Guo Y, Yarbrough WG, Ivanov SV. Expression profiling of clinical specimens supports the existence of neural progenitor-like stem cells in basal breast cancers. Clin Breast Cancer. 2017;17(4):298–306 e7.
116. Harbhajanka A, Chahar S, Miskimen K, Silverman P, Harris L, Williams N, et al. Clinicopathological, immunohistochemical and molecular correlation of neural crest transcription factor SOX10 expression in triple-negative breast carcinoma. Hum Pathol. 2018;80:163–9.
117. Laurent E, Begueret H, Bonhomme B, Veillon R, Thumerel M, Velasco V, et al. SOX10, GATA3, GCDFP15, androgen receptor, and mammaglobin for the differential diagnosis between triple-negative breast cancer and TTF1-negative lung adenocarcinoma. Am J Surg Pathol. 2019;43(3):293-302.
118. Bilir B, Osunkoya AO, Wiles WG, Sannigrahi S, Lefebvre V, Metzger D, et al. SOX4 is essential for prostate tumorigenesis initiated by PTEN Ablation. Cancer Res. 2016;76(5):1112–21.
119. Bilir B, Kucuk O, Moreno CS. Wnt signaling blockage inhibits cell proliferation and migration, and induces apoptosis in triple-negative breast cancer cells. J Transl Med. 2013;11:280.
120. Tiwari N, Tiwari VK, Waldmeier L, Balwierz PJ, Arnold P, Pachkov M, et al. Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. Cancer Cell. 2013;23(6):768–83.
121. Vervoort SJ, Lourenco AR, Tufegdzic Vidakovic A, Mocholi E, Sandoval JL, Rueda OM, et al. SOX4 can redirect TGF-beta-mediated SMAD3-transcriptional output in a context-dependent manner to promote tumorigenesis. Nucleic Acids Res. 2018;46(18):9578–90.
122. Lee AK, Ahn SG, Yoon JH, Kim SA. Sox4 stimulates ss-catenin activity through induction of CK2. Oncol Rep. 2011;25(2):559–65.
123. Lopez-Knowles E, O'Toole SA, McNeil CM, Millar EK, Qiu MR, Crea P, et al. PI3K pathway activation in breast cancer is associated with the basal-like phenotype and cancer-specific mortality. Int J Cancer. 2010;126(5):1121–31.
124. Zhang Y, Kwok-Shing Ng P, Kucherlapati M, Chen F, Liu Y, Tsang YH, et al. A pan-cancer proteogenomic atlas of PI3K/AKT/mTOR pathway alterations. Cancer Cell. 2017;31(6): 820–32.e3.
125. Fruman DA, Rommel C. PI3K and cancer: lessons, challenges and opportunities. Nat Rev Drug Discov. 2014;13(2):140–56.
126. Hallstrom TC, Mori S, Nevins JR. An E2F1-dependent gene expression program that determines the balance between proliferation and cell death. Cancer Cell. 2008;13:11–22.
127. Rodon J, Dienstmann R, Serra V, Tabernero J. Development of PI3K inhibitors: lessons learned from early clinical trials. Nat Rev Clin Oncol. 2013;10(3):143–53.
128. Wong KK, Engelman JA, Cantley LC. Targeting the PI3K signaling pathway in cancer. Curr Opin Genet Dev. 2010;20(1):87–90.
129. Ramezani-Rad P, Geng H, Hurtz C, Chan LN, Chen Z, Jumaa H, et al. SOX4 enables oncogenic survival signals in acute lymphoblastic leukemia. Blood. 2013;121(1):148–55.
130. Polakis P. Wnt signaling in cancer. Cold Spring Harb Perspect Biol. 2012;4(5):a008052.
131. Khramtsov AI, Khramtsova GF, Tretiakova M, Huo D, Olopade OI, Goss KH. Wnt/beta-catenin pathway activation is enriched in basal-like breast cancers and predicts poor outcome. Am J Pathol. 2010;176(6):2911–20.
132. Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, et al. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. J Biol Chem. 2008;283(26):17969–78.
133. Ye X, Wu F, Wu C, Wang P, Jung K, Gopal K, et al. beta-Catenin, a Sox2 binding partner, regulates the DNA binding and transcriptional activity of Sox2 in breast cancer cells. Cell Signal. 2014;26(3):492–501.
134. Liu K, Xie F, Gao A, Zhang R, Zhang L, Xiao Z, et al. SOX2 regulates multiple malignant processes of breast cancer

development through the SOX2/miR-181a-5p, miR-30e-5p/TUSC3 axis. Mol Cancer. 2017;16(1):62.
135. Wang J, Zeng H, Li H, Chen T, Wang L, Zhang K, et al. MicroRNA-101 inhibits growth, proliferation and migration and induces apoptosis of breast cancer cells by targeting sex-determining region Y-Box 2. Cell Physiol Biochem. 2017;43(2):717–32.
136. Zhu YT, Jia Y, Hu L, Qi C, Prasad MK, McCallion AS, et al. Peroxisome-proliferator-activated receptor-binding protein (PBP) is essential for the growth of active Notch4-immortalized mammary epithelial cells by activating SOX10 expression. Biochem J. 2009;425(2):435–44.
137. Cao Q, Chen X, Wu X, Liao R, Huang P, Tan Y, et al. Inhibition of UGT8 suppresses basal-like breast cancer progression by attenuating sulfatide-alphaVbeta5 axis. J Exp Med. 2018;215(6):1679–92.
138. Wang QY, Zhou CX, Zhan MN, Tang J, Wang CL, Ma CN, et al. MiR-133b targets Sox9 to control pathogenesis and metastasis of breast cancer. Cell Death Dis. 2018;9(7):752.
139. Zhao Y, Pang W, Yang N, Hao L, Wang L. MicroRNA-511 inhibits malignant behaviors of breast cancer by directly targeting SOX9 and regulating the PI3K/Akt pathway. Int J Oncol. 2018;53(6):2715–26.
140. Chen X, Fu Y, Xu H, Teng P, Xie Q, Zhang Y, et al. SOX5 predicts poor prognosis in lung adenocarcinoma and promotes tumor metastasis through epithelial-mesenchymal transition. Oncotarget. 2018;9(13):10891–904.
141. Zhang D, Liu S. SOX5 promotes epithelial-mesenchymal transition in osteosarcoma via regulation of Snail. J BUON. 2017;22(1):258–64.
142. Si C, Yu Q, Yao Y. Effect of miR-146a-5p on proliferation and metastasis of triple-negative breast cancer via regulation of SOX5. Exp Ther Med. 2018;15(5):4515–21.
143. Young N, Hahn CN, Poh A, Dong C, Wilhelm D, Olsson J, et al. Effect of disrupted SOX18 transcription factor function on tumor growth, vascularization, and endothelial development. J Natl Cancer Inst. 2006;98(15):1060–7.
144. Zhang J, Ma Y, Wang S, Chen F, Gu Y. Suppression of SOX18 by siRNA inhibits cell growth and invasion of breast cancer cells. Oncol Rep. 2016;35(6):3721–7.
145. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119(6):1420–8.
146. Li X, Xu Y, Chen Y, Chen S, Jia X, Sun T, et al. SOX2 promotes tumor metastasis by stimulating epithelial-to-mesenchymal transition via regulation of WNT/beta-catenin signal network. Cancer Lett. 2013;336(2):379–89.
147. Pang Y, Liu J, Li X, Xiao G, Wang H, Yang G, et al. MYC and DNMT3A-mediated DNA methylation represses microRNA-200b in triple negative breast cancer. J Cell Mol Med. 2018;22(12):6262–74.
148. Vervoort SJ, Lourenco AR, van Boxtel R, Coffer PJ. SOX4 mediates TGF-beta-induced expression of mesenchymal markers during mammary cell epithelial to mesenchymal transition. PLoS One. 2013;8(1):e53238.
149. Heldin CH, Vanlandewijck M, Moustakas A. Regulation of EMT by TGFbeta in cancer. FEBS Lett. 2012;586(14):1959–70.
150. Jafarnejad SM, Wani AA, Martinka M, Li G. Prognostic significance of Sox4 expression in human cutaneous melanoma and its role in cell migration and invasion. Am J Pathol. 2010;177(6):2741–52.
151. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, et al. Endogenous human microRNAs that suppress breast cancer metastasis. Nature. 2008;451(7175):147–52.
152. Xi J, Feng J, Zeng S. Long noncoding RNA lncBRM facilitates the proliferation, migration and invasion of ovarian cancer cells via upregulation of Sox4. Am J Cancer Res. 2017;7(11):2180–9.
153. Yang M, Wang J, Wang L, Shen C, Su B, Qi M, et al. Estrogen induces androgen-repressed SOX4 expression to promote progression of prostate cancer cells. Prostate. 2015;75(13):1363–75.
154. Zhou Y, Wang X, Huang Y, Chen Y, Zhao G, Yao Q, et al. Down-regulated SOX4 expression suppresses cell proliferation, metastasis and induces apoptosis in Xuanwei female lung cancer patients. J Cell Biochem. 2015;116(6):1007–18.
155. Lee H, Goodarzi H, Tavazoie SF, Alarcon CR. TMEM2 is a SOX4-regulated gene that mediates metastatic migration and invasion in breast cancer. Cancer Res. 2016;76(17):4994–5005.
156. Liu S, Patel SH, Ginestier C, Ibarra I, Martin-Trevino R, Bai S, et al. MicroRNA93 regulates proliferation and differentiation of normal and malignant breast stem cells. PLoS Genet. 2012;8(6):e1002751.
157. Bai JW, Wang X, Zhang YF, Yao GD, Liu H. MicroRNA-320 inhibits cell proliferation and invasion in breast cancer cells by targeting SOX4. Oncol Lett. 2017;14(6):7145–52.
158. Hanieh H. Aryl hydrocarbon receptor-microRNA-212/132 axis in human breast cancer suppresses metastasis by targeting SOX4. Mol Cancer. 2015;14:172.
159. Jin Y, Zhao M, Xie Q, Zhang H, Wang Q, Ma Q. MicroRNA-338-3p functions as tumor suppressor in breast cancer by targeting SOX4. Int J Oncol. 2015;47(4):1594–602.
160. Wang N, Liu W, Zheng Y, Wang S, Yang B, Li M, et al. CXCL1 derived from tumor-associated macrophages promotes breast cancer metastasis via activating NF-kappaB/SOX4 signaling. Cell Death Dis. 2018;9(9):880.
161. Yang F, Shen Y, Zhang W, Jin J, Huang D, Fang H, et al. An androgen receptor negatively induced long non-coding RNA ARNILA binding to miR-204 promotes the invasion and metastasis of triple-negative breast cancer. Cell Death Differ. 2018;25(12):2209-220.
162. Hu J, Tian J, Zhu S, Sun L, Yu J, Tian H, et al. Sox5 contributes to prostate cancer metastasis and is a master regulator of TGF-beta-induced epithelial mesenchymal transition through controlling Twist1 expression. Br J Cancer. 2018;118(1):88–97.
163. Renjie W, Haiqian L. MiR-132, miR-15a and miR-16 synergistically inhibit pituitary tumor cell proliferation, invasion and migration by targeting Sox5. Cancer Lett. 2015;356(2 Pt B):568–78.
164. Wang D, Han S, Wang X, Peng R, Li X. SOX5 promotes epithelial-mesenchymal transition and cell invasion via regulation of Twist1 in hepatocellular carcinoma. Med Oncol. 2015;32(2):461.
165. Yang B, Zhang W, Sun D, Wei X, Ding Y, Ma Y, et al. Downregulation of miR-139-5p promotes prostate cancer progression through regulation of SOX5. Biomed Pharmacother. 2019;109:2128–35.
166. Zhang YJ, Xu F, Zhang YJ, Li HB, Han JC, Li L. miR-206 inhibits non small cell lung cancer cell proliferation and invasion by targeting SOX9. Int J Clin Exp Med. 2015;8(6):9107–13.
167. Narasimhan K, Pillay S, Bin Ahmad NR, Bikadi Z, Hazai E, Yan L, et al. Identification of a polyoxometalate inhibitor of the DNA binding activity of Sox2. ACS Chem Biol. 2011;6(6):573–81.
168. Chen X, Zheng Q, Li W, Lu Y, Ni Y, Ma L, et al. SOX5 induces lung adenocarcinoma angiogenesis by inducing the expression of VEGF through STAT3 signaling. Onco Targets Ther. 2018;11:5733–41.
169. Yang H, Lee S, Lee S, Kim K, Yang Y, Kim JH, et al. Sox17 promotes tumor angiogenesis and destabilizes tumor vessels in mice. J Clin Invest. 2013;123(1):418–31.
170. Bojang P Jr, Ramos KS. The promise and failures of epigenetic therapies for cancer treatment. Cancer Treat Rev. 2014;40(1):153–69.

171. Mund C, Lyko F. Epigenetic cancer therapy: proof of concept and remaining challenges. Bioessays. 2010;32(11):949–57.
172. Miranda TB, Cortez CC, Yoo CB, Liang G, Abe M, Kelly TK, et al. DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. Mol Cancer Ther. 2009;8(6):1579–88.
173. Tan J, Yang X, Zhuang L, Jiang X, Chen W, Lee PL, et al. Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. Genes Dev. 2007;21(9):1050–63.
174. Hur W, Rhim H, Jung CK, Kim JD, Bae SH, Jang JW, et al. SOX4 overexpression regulates the p53-mediated apoptosis in hepatocellular carcinoma: clinical implication and functional analysis in vitro. Carcinogenesis. 2010;31(7):1298–307.
175. Matheu A, Collado M, Wise C, Manterola L, Cekaite L, Tye AJ, et al. Oncogenicity of the developmental transcription factor Sox9. Cancer Res. 2012;72(5):1301–15.
176. Zhu Y, Li Y, Jun Wei JW, Liu X. The role of Sox genes in lung morphogenesis and cancer. Int J Mol Sci. 2012;13(12):15767–83.
177. Andreucci E, Pietrobono S, Peppicelli S, Ruzzolini J, Bianchini F, Biagioni A, et al. SOX2 as a novel contributor of oxidative metabolism in melanoma cells. Cell Commun Signal. 2018;16(1):87.
178. Bhattaram P, Muschler G, Wixler V, Lefebvre V. Inflammatory cytokines stabilize SOXC transcription factors to mediate the transformation of fibroblast-like synoviocytes in arthritic disease. Arthritis Rheum. 2018;70(3):371–82.