Transcription Factor Sox4 as a Potential Player in Mammary Gland Involution

Hirohisa Shibayama, Tomomi Yamamoto, Kenzi Oshima,

Tsukasa Matsuda, and Daita Nadano

Department of Applied Biosciences, Graduate School of Bioagricultural Sciences,

Nagoya University, Nagoya 464-8601, Japan

Running title: Sox4 in mammary gland involution

Address correspondence to:

Daita Nadano, Ph. D.

Department of Applied Biosciences

Graduate School of Bioagricultural Sciences

Nagoya University

Furo-cho, Chikusa, Nagoya 464-8601

Japan

E-mail: nadano@agr.nagoya-u.ac.jp

Abstract

Mammary gland involution is a regressive process for the gland to return to its pre-pregnancy state after lactation and comprises an initial reversible and second remodeling stage. Although many genes and the multiple expression profiles of their mRNAs have been found in this process, the mechanisms controlling the profiles are largely unknown. Here, we identified and analyzed transcription factor Sox4 in mammary gland involution. Elevated expression of Sox4 gene in the first stage (48 h after weaning) was observed at the mRNA and protein levels in the mouse mammary gland. Immunohistochemistry of the involuting gland indicated that Sox4 was located in the nuclei of epithelial cells. Nuclear Sox4 was also detected in the second stage, but unlikely to be involved in cell death, one of the characteristic events of involution. To clarify the functional roles of Sox4 in involution, we introduced a model, including a normal mammary epithelial cell line, for finding candidate target genes of this transcription factor and examined its effect on tenascin C mRNA expression.

Key words: mammalian development, mammary gland involution, gene expression, transcriptional control, Sox4, cancer

Introduction

The mammary gland is a unique tissue that can provide newborns with milk. Pregnancy triggers the proliferation and differentiation of the mammary gland, resulting in the expansion of the lobuloalveolar units (Macias & Hinck 2012). The content of adipocytes, which occupy 97% of the nulliparous mouse mammary gland, decreases to 11% in lactation (McCready *et al.* 2014). Milk, including various proteins, lipids, and other nutrients, is secreted after delivery from the fully differentiated epithelial cells of the secretory alveoli. After weaning, the gland shifts to a regressive phase, involution, and is remodeled back to a state resembling that of the adult nulliparous female. Thus, the mammary gland is also unique to exhibit this dynamic developmental cycle of pregnancy, lactation, and involution (Macias & Hinck 2012).

The mechanisms underpinning lactation have long and extensively investigated because they have been not only of biological interest, but also useful for dairy applications. Compared to this, the studies of involution have lagged behind. Originally according to morphological observations, involution has been divided into two stages (Lund *et al.* 1996; Wiseman & Werb 2002; Nakamura *et al.* 2006; Jena *et al.* 2019). In the first stage (for the first 48-72 h after weaning in mice), no major morphological changes occur in the gland, whereas milk accumulation and shedding of dying cells are observed in the alveolar lumen. Re-suckling by pups can re-initiate lactation in this reversible stage. The second stage has been reported to commence at about 48 h post weaning in mice. Compositional and structural changes in the extracellular matrix and apoptosis of epithelial cells become evident, which is followed by irreversible tissue remodeling including alveoli collapse and adipocyte refilling. The molecular mechanisms underlying these phenomena have been reported in this decade, which include the pivotal role of the transcription factor Stat3 in coordinating the involution process (Hughes & Watson 2018), both in terms of lysosomal-mediated cell death (Kreuzaler *et al.* 2011; Sargeant *et al.* 2014), and in terms of modulation of the

microenvironment (Hughes *et al.* 2012). Involution has been linked to breast cancer progression. For example, the involuting mammary gland has similar characteristics to tumor promotional microenvironment, and extracellular matrix isolated from involuting mammary glands was found to facilitate invasiveness of mammary tumor cells (Lyons *et al.* 2011; Fornetti *et al.* 2014). Further molecular analysis of this phase has been expected for the prevention and treatment of breast cancer.

Comprehensive transcriptome and proteome analyses have identified a lot of genes expressed in mammary gland involution at the mRNA and protein levels, respectively (Rudolph *et al.* 2003; Clarkson *et al.* 2004; Davies *et al.* 2006). DNA microarray analysis has defined multiple expression profiles of the genes in this phase (Stein *et al.* 2007). Although these excellent data have explored avenues for a deeper understanding of involution biology, it remains unresolved how these profiles are strictly controlled by regulatory molecules including transcription factors. In the present study, we report that Sox4 is potentially included in transcriptional control in involution.

Materials and Methods

Mice, cells, and nuclear isolation

Animal experiments of this study were approved by the Committee on Animal Experiments of Nagoya University. Mouse mammary glands were collected according to our previous report (Nakamura *et al.* 2006). Synchronous involution was induced at day 10 of lactation (L10) by the removal of suckling pups. Ovary and thymus tissues were collected from adult mice (about seven weeks old). For nuclei isolation, the mammary gland was minced, transferred into a Potter-type homogenizer, and homogenized in 5 mM Tris-HCl, pH 7.4, containing 0.25 M sucrose and 3.3 mM $CaCl_2$. The homogenate was centrifuged at 600g for 10 min at 4ºC, and nuclei were collected as pellet. Human HEK293T and MCF-7 cell lines were cultured as described (Muto *et al.* 2018). Mouse mammary cell line HC11 (Ball *et al.* 1988) was kindly provided by Dr. B. Groner (Georg Speyer Haus, Institute for Biomedical Research, Frankfurt am Main, Germany) and maintained at 37ºC under 5% CO_2 in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum, 10 ng/ ml epidermal growth factor (Nakatani *et al.* 2010).

mRNA analysis

Total cellular RNA was isolated from mouse tissues and HC11 cells by using Trizol reagent (Invitrogen, Carlsbad, CA, USA). Reverse transcription (RT) was performed by using SuperScript II (Invitrogen) (Nakamura *et al.* 2006). To assess the mRNA expression of mouse Sox4, partial cDNA of mouse Sox4 was amplified by semi-quantitative polymerase chain reaction (PCR) by using the RT product, Taq DNA polymerase (Takara, Japan), and the following primer pair: 5'-AGTGAAGCGCGTCTACCTGT-3' and 5'-TCAGACTCCGGCCATCGG-3'. PCR

cycling began with template denaturation at 94ºC for 4 min, then 30 cycles of 94ºC for 1 min, 62ºC for 1 min, and 72ºC for 1 min. Partial cDNA of mouse tenascin C was amplified in the same manner, excepting that cycle number and annealing temperature were altered to 35 and 60ºC, respectively, and that the following primer pair was used: 5'-ACAAGACCATGGGGAGTCTG-3' and 5'-TGGCTGAGTCTGTGTCCTTG-3'. Mouse glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA was amplified as described (Muto *et al.* 2018). Amplified DNAs were subjected to electrophoresis on 2% agarose gels or 8% polyacrylamide gels and stained with ethidium bromide. The intensities of amplified tenascin C and GAPDH cDNA bands on polyacrylamide gels were quantified by using the CS analyzer software (Atto, Tokyo, Japan) to calculate relative transcript levels in HC11 cells. The relative tenascin C transcript level was normalized to that of GAPDH. Student's t-test was used to evaluate significance.

Sox4 expression vectors

The 5'-terminal fragment (the initiation ATG site to the unique SalI site) of mouse Sox4 cDNA was amplified by PCR using the mammary gland cDNA. A cDNA clone (no. AK028989) including the 3'-terminal part (the SalI site to the termination TGA site) of mouse Sox4 cDNA was obtained from RIKEN (Yokohama, Japan). These fragments were inserted in-frame into the pCMV-Tag2B (Agilent Technologies, Santa Clara, CA, USA) to express N-terminally FLAG-tagged full-length Sox4. The vector for the expression of the C-terminally truncated form of FLAG-tagged mouse Sox4 (residues 1-408) was prepared by PCR (Imai *et al.* 1991) using the full-length Sox4 expression vector and the following primer pair: 5'-TGAAGGGAGCGCGGC-3' and 5'-GGAGCCTGAGCCGGGTTC-3'. The cDNA encoding full-length Sox4 was cloned into pEGFP-C1 (Clontech, Mountain View, CA, USA) to express Sox4 N-terminally fused with green fluorescent protein (GFP).

Vector transfection

PEI-Max (100 mg; Polysciences, Warrington, PA, USA) was dissolved in about 90 ml sterilized water. After pH adjustment to 7.0 with NaOH, the solution was made up to 100 ml (final concentration, 1.0 mg/ml) with water, filtrated (0.2 μm), aliquoted, and stored at -30°C until use. For transient transfection in HC11 cells, cells were seeded at 3.0 x 10^5/ dish in a 35-mm culture dish and incubated for 20 h. Vector DNA (3-30 μg) was mixed with 100 μl 1x phosphate-buffered saline in a 1.5-ml tube and vortexed. The PEI-Max solution (8-90 μl) was mixed with 100 μl 1x phosphate-buffered saline in another 1.5-ml tube and vortexed. These solutions were then mixed by vortexing, kept for 15 min at room temperature, and added dropwise to the cell plate. After gentle swirling, the cells were incubated at 37°C under 5% CO_2. The medium was removed after 5 h. The cells were refed with fresh, 37°C-prewarmed medium and incubated for 19 h.

Transient transfection in HEK293T and MCF-7 cells was described previously (Muto *et al.* 2018). To obtain stable transfectants, pEGFP-C1 or the expression vector of GFP-fused Sox4 was introduced into MCF-7 cells and cultured in the presence of 1 mg/ml G418 (Calbiochem, Merck Biosciences, Darmstadt, Germany). About 3 weeks after transfection, G418-resistant colonies were observed under a fluorescence microscope. Each colony exhibiting GFP fluorescence was picked up, cultured, and subjected to immunoblotting for final selection.

Protein analyses

For the preparation of anti-mouse Sox4 antibody, a synthetic peptide, corresponding to residues 287-308 of mouse Sox4 (Figure 1A), was conjugated to keyhole limpet hemocyanin by using Imject maleimide activated mcKLH (Piece, Thermo Fisher Scientific, Waltham, MA, USA) and injected into rabbits. The antibody

was affinity-purified using the antigenic peptide (Miyoshi *et al.* 2007). Rabbit polyclonal antibody against MFG-E8, which was strongly expressed in the involuting mammary gland, was described previously (Nakatani *et al.* 2006). Mouse antibodies against a FLAG tag (M2) and α-tubulin were purchased from Sigma-Aldrich (St. Louis, MO, USA). Mouse anti-fibrillarin antibody was purchased from EnCor Biotechnology (Gainesville, FL, USA). Rabbit anti-GFP antibody was purchased from Medical & Biological Laboratories (Nagoya, Japan). Guinea pig polyclonal antibody recognizing keratins K8 and K18 (GP11) was purchased from Progen (Heidelberg, Germany).

Immunoblotting and antibody preabsorption with the antigenic peptide were performed as described (Miyoshi *et al.* 2007; Sugihara *et al.* 2013). To evaluate the relative expression levels of the full-length and mutant forms of FLAG-tagged Sox4 in HC11 cells, the band intensities in immunoblots (probed with anti-FLAG antibody) were quantified by using the CS analyzer software. Immunohistochemistry using frozen sections and a confocal laser scan microscope was described previously (Sugihara *et al.* 2013). Nuclei were stained with TOTO-3 (Molecular Probes, Thermo Fisher Scientific) or propidium iodide (Sigma-Aldrich). Apoptotic cells were detected by the terminal deoxynucleotidyl transferase-mediated dUTP-biotin nick end labeling (TUNEL) method (In situ cell death detection kit, Roche, Sigma-Aldrich).

Results

Sharp increase of Sox4 expression in the mammary gland after pup weaning

We have previously surveyed genes whose mRNA expression was upregulated 48 h after weaning (or at day 2 of involution, In2) in the mouse mammary gland (Nakamura *et al.* 2006). This time point has been recognized to be just before or at the beginning of the second stage of involution. Hence, gene expression at In2 has been assumed to be important for a smooth transition to the second stage. Sox4 was thus identified as a candidate gene in our survey. RT-PCR confirmed that the expression of Sox4 was elevated at In2, as compared to its expression in mid-lactation (L10) (Figure 1B). The mRNA elevation was detected 24 h after weaning (at In1), and the previous data showing Sox mRNA expression in the ovary and thymus (van de Wetering *et al.* 1993) were confirmed under these RT-PCR conditions (Figure 1C).

Sox4 is a transcription factor possessing the N-terminal DNA binding domain (HMG domain) and the C-terminal transactivation domain (Figure 1A), and is one of about 20 members of the Sox family (Wegner 1999; Vervoort *et al.* 2013). These members play vital roles in normal development in mammals, and the importance of Sox4 has been reported in lymphocyte and neuronal differentiations (Penzo-Mendez 2010). Whereas Sox4 mRNA has been reported be expressed at very low abundance in the mammary glands of pubertal and non-pregnant adult mice (Hunt & Clarke 1999), the relationship between Sox4 and the developmental cycle of the gland (Macias & Hinck 2012; Jena *et al.* 2019) has not been documented. Hence, we analyzed this Sox4 upregulation further.

The upregulation of Sox4 was then examined at the protein level. We tried to use commercially available antibodies against human Sox4, which showed a poor cross-reactivity against mouse Sox4 (data not shown). We prepared new antibody against mouse Sox4 by immunizing rabbits with the synthetic peptide (Figure 1A). This

antibody showed a specific reactivity with FLAG-tagged mouse Sox4 expressed in cells (Figure 2A). Immunoblotting with this antibody indicated high expression of Sox4 in the mouse mammary gland at In2 (Figure 2B). Pretreatment of the antibody with the antigenic peptide abolished the detection, excluding the possibility of false reactivity of the antibody IgG (Figure 2C). The expression of Sox4 at In1 (Figure 1C) was confirmed at the protein level (Figure 2D). Expression of Sox4 in adult tissues is restricted, including pancreatic islet cells, gonads, and thymus (Penzo-Mendez 2010). The expression levels of Sox4 were comparable in the mammary gland (early involution), ovary, and thymus (Figures 1C and 2D). These results indicate a new localization of Sox4 in normal adult mammals.

Sox4 was located in the nuclei of mammary epithelial cells at In2

In lactation, the mammary gland is filled with the lobuloalveolar units including milk-secreting epithelial cells, and the general structure of the mammary gland is maintained for the first 2 days after weaning (Wiseman & Werb 2002). It was therefore supposed that Sox4 at In2 was expressed in the epithelial cells. Similar to other Sox proteins (Malki *et al.* 2010), Sox4 has been reported to display nucleocytoplasmic shuttling properties within the cell (Beekman *et al.* 2012). We investigated tissue and subcellular localization of Sox4 by immunohistochemistry with our anti-Sox4 antibody. In the mammary gland at In2, Sox4 was detected in the monolayer cells lining the acini and localized mainly in the nuclei within these monolayer cells (Figure 3A; see also the control in Figure 3B as a magnified view). The signals detected with the same antibody were vague at L10, which was consistent with the relatively low expression of Sox4 at the time (Figures 1B and 2B). The localization of Sox4 in nuclei, where, needless to say, transcription occurs, at In2 was confirmed by the two different experiments. First, these strong nuclear signals were lost in the peptide blocking experiments (Figure 3B). Second, the putative Sox4 band was detected by

immunoblotting of the nuclei isolated from the mammary gland at In2 (Figure 3C). Co-localization of Sox4 and keratins has indicated that Sox4 is expressed in epithelial cells (Figure 4A). Overall, these results do not seem inconsistent with the assumption that Sox4 functions as a transcription factor in involution.

Nuclear Sox4 was observed in the second stage of involution

Nuclear Sox4 was found at In2, which is a transition into the second stage of involution (Macias & Hinck 2012). It is possible that a gene expressed at this time point participates in the irreversible execution of involution in the second stage. According to this assumption, expression of the gene is likely to be retained after In2. Or, if the gene functions mainly in the first stage, the expression would decline rapidly after the time. To distinguish these two possibilities, we examined Sox4 in the second stage of involution.

In the mouse mammary gland, characteristic phenomena, including gland remodeling, in the second stage are widely observed at day 4 of involution (In4), and the majority of the phenomena is complete from In5 to In6 (Atabai *et al.* 2007; Macias & Hinck 2012). Immunohistochemistry of tissue sections at In4 indicated that nuclear Sox4 was still clearly observed (Figure 4B), supporting the former possibility. Two distinct types of cell death have been reported in mammary gland involution: lysosomal-mediated, non-apoptotic, cell death in the first stage (Kreuzaler *et al.* 2011; Sargeant *et al.* 2014; Hughes & Watson 2018) and proteinase-dependent apoptosis in the second stage (Lund *et al.* 1996; Wiseman & Werb 2002; Jena *et al.* 2019). Fragmentation of nuclear DNA was observed in the sections of In4 by using TUNEL assay (Figure 4B), indicating apoptosis of unnecessary secretory epithelial cells, and the observed signals did not show obvious colocalization with those of Sox4. Although Sox4 has been reported to execute apoptosis in a subset of tumors (Pan *et al.* 2009; Jafarnejad *et al.* 2013), Sox4 was unlikely be a direct driver of apoptosis in mammary

gland involution.

HC11 cells expressing Sox4 might be a model to analyze its target genes in mammary gland involution

Our data did not show direct regulation of apoptosis by Sox4 in involution. Although apoptosis is one of the characteristic events of involution, this phase is not a simple removal of the majority (90%) of mammary epithelial cells, but rather involves remodeling of the mammary gland for next pregnancy (Wiseman & Werb 2002; Jena *et al.* 2019). The sustained Sox4 expression in involution has led to the hypothesis that there are many Sox4 targets in mammary epithelial cells and that they are involved in the large dynamic changes which occur in the period at the whole tissue level.

For comprehensive identification of the candidate target genes, a simple, scalable method is suitable, such as the use of culture cells transiently transfected with a Sox4 expression vector. Since most of mammary epithelial cells finally die during involution, it has been considered to be difficult to obtain a convenient cell model for involuting epithelial cells, although KIM-2 murine mammary epithelial cells have been used extensively to model involution (Gordon *et al.* 2000; Hughes *et al.* 2016). We adopted HC11, a cell line established from the normal epithelial cells of the mouse mammary grand (Ball *et al.* 1988). This cell line possesses the properties of mammary epithelial cells, including hormonal induction of milk protein synthesis (Ball *et al.* 1988), and constitutively expresses Stat3 (Philp *et al.* 1996), which has key roles in involution (Hughes & Watson 2018).

Because the transfection efficiency of expression vectors in HC11 was very low under our initial conditions (Figure 5A, top images), the conditions were examined by changing the amounts of a GFP expression vector and polyethyleneimine (PEI-Max), which is an efficient, inexpensive transfection reagent and applicable in vitro and in vivo (Neuberg & Kichler 2014). The efficiency was obviously improved under the

optimized conditions (3 μg vector DNA and 60 μg PEI-Max per 35-mm dish), as assessed by immunoblotting (Figure 5B) and fluorescence microscopy (Figure 5A, middle images).

As part of our approach, we tested the effect of Sox4 on tenascin C expression. Tenascin C is an extracellular matrix protein and a member of the tenascin family. This family proteins have been shown to be involved in tissue remodeling (Jones & Jones 2000), which Sox4 can promote (Tiwari *et al.* 2013; Vervoort *et al.* 2013). Tenascin C has been reported to be secreted around epithelial basement lines at In2 in the mouse mammary gland (Jones *et al.* 1995; Fornetti *et al.* 2014). Using our model, we asked whether the elevated expression of Sox4 and tenascin C in involution was a simple coincidence. The expression of FLAG-tagged Sox4 was observed under the optimized transfection conditions in HC11 (Figure 6A). Upregulation of tenascin C mRNA was observed with overexpression of Sox4 (Figure 6B, left graph), suggesting a causal relationship between the two events. Although the mRNA increase was 1.8-fold on average, this low induction might be connected to the intracellular instability of full-length Sox4 protein (see Discussion). We also used the deletion mutant of mouse Sox4, which lacks the very C-terminal part of the protein (CT in Figure 1A) and hardly activates target genes (Dy *et al.* 2008). The upregulation of tenascin C by FLAG-tagged Sox4 was markedly compromised by this deletion (Figure 6B, right graph), which did not seem to support the possibility that the upregulation (Figure 6B, left graph) was simply due to aberrant transactivation including overexpressed Sox4.

Discussion

To date, several transcription factors have been found to function in mammary glad involution. Stat3 plays multiple roles in the first and second stages of involution (Hughes & Watson 2018). p53-null mice have demonstrated the importance of the protein in this regressive phase including p53-dependent apoptosis (Jerry *et al.* 2002; Wiseman & Werb 2002). Moreover, mRNA stability has been reported to be controlled by microRNAs (Jena *et al.* 2019). Nonetheless, much more has yet to be discovered to understand the regulation of the complex mRNA profiles in involution (Rudolph *et al.* 2003; Clarkson *et al.* 2004; Stein *et al.* 2007). Here, we identified Sox4 as a potential player in this process. At the tissue level, the expression of Sox4 in the involuting mammary gland was relatively high (Figures 1C and 2D). This may be related to the fact that most of the gland in early involution is occupied by the lobuloalveolar units including epithelial cells (Wiseman & Werb 2002), where Sox4 was found to be expressed (Figures 3A and 4A). Sox4 has been reported to control transcription by interacting with other proteins (Wilson & Koopman 2002; Kondoh & Kamachi 2010). Specifically, Sox4 and p53 have been shown to cooperate to proceed apoptosis in DNA damage response-associated cancer (Pan *et al.* 2009). Although our data (Figure 4B) have suggested that such cooperative promotion of apoptosis does not work in involution, this is not surprising since Sox4, as other Sox members do (Wegner 1999), acts in a cell-type and cell-context dependent manner (Jafarnejad *et al.* 2013; Vervoort *et al.* 2013). Stat3-mediated Sox4 expression has been found in liver cancer (Chen *et al.* 2016). It seems intriguing to address whether these two factors are linked in a similar or different manner in involution. The target genes of Sox4 in human cancers have been surveyed by methods including transient overexpression, RNA interference, and chromatin immunoprecipitation-chip analysis, and numerous candidates (> 100 per dataset) were identified (Vervoort *et al.* 2013). For instance, a systematic survey of Sox4 targets in prostate cancer cells has revealed 283 high-confidence candidates

including tenascin C (Scharer *et al.* 2009). Compared to these findings from cancer, Sox4 targets in normal development remain largely unsolved. Surveys of Sox4 targets in involution, including the use of our model (Figure 6), might provide a new insight into the Sox4 transcriptional network.

Full-length Sox4 protein has been reported to be very unstable in cultured cells (half-life, < 1 h) (Beekman *et al.* 2012). This instability was dependent on the C-terminal 33 residues of Sox4 (Beekman *et al.* 2012) (CT in Figure 1A), which was also essential for transactivation of target genes (Dy *et al.* 2008). Our results confirmed the importance of this region for Sox4 protein stability by showing that, under the same transfection conditions, the expression level of full-length Sox4 was much lower than that of Sox4 lacking the region (Figure 6A). The importance of the C-terminal region seemed to explain why we were unable to establish cells stably expressing full-length Sox4 (Supplementary Figure S1). The nuclear expression of Sox4 was detected at In2 and In4 (Figures 3A and 4B). This suggests stable Sox4 expression for at least 2 days and lets us speculate a mechanism actively controlling the level of Sox4 protein in the mammary gland. Sox4 was shown to be degraded by the proteasome in a polyubiquitin-independent manner (Beekman *et al.* 2012). However, the mechanistic details in rapid Sox4 degradation have been unresolved. Coupled with the multifaceted roles of Sox4 in carcinogenesis (Jafarnejad *et al.* 2013; Vervoort *et al.* 2013), further analysis of Sox4 expression in the involuting mammary gland is expected to pursue the possibility of controlling and manipulating Sox4 stability in vivo.

Acknowledgements

We thank Satoshi Gamano and Yumejichi Fujita (Nagoya University) for technical supports and helpful discussions. This work was supported in part by the Ito Foundation (Tokyo, Japan).

Disclosure Statement

No competing financial interests exist.

References

Atabai, K., Sheppard, D. & Werb, Z. (2007) Roles of the innate immune system in mammary gland remodeling during involution. J. Mammary Gland Biol. Neoplasia **12**, 37-45.

Ball, R.K., Friis, R.R., Schoenenberger, C.A., Doppler, W. & Groner, B. (1988) Prolactin regulation of beta-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line. EMBO J. **7**, 2089-2095.

Beekman, J.M., Vervoort, S.J., Dekkers, F., van Vessem, M.E., Vendelbosch, S., Brugulat-Panes, A., van Loosdregt, J., Braat, A.K. & Coffer, P.J. (2012) Syntenin-mediated regulation of Sox4 proteasomal degradation modulates transcriptional output. Oncogene **31**, 2668-2679.

Chen, Z.Z., Huang, L., Wu, Y.H., Zhai, W.J., Zhu, P.P. & Gao, Y.F. (2016) LncSox4 promotes the self-renewal of liver tumour-initiating cells through Stat3-mediated Sox4 expression. Nat. Commun. **7**, 12598.

Clarkson, R.W.E., Wayland, M.T., Lee, J., Freeman, T. & Watson, C.J. (2004) Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. Breast Cancer Res. **6**, R92-R109.

Davies, C.R., Morris, J.S., Griffiths, M.R., Page, M.J., Pitt, A., Stein, T. & Gusterson, B.A. (2006) Proteomic analysis of the mouse mammary gland is a powerful tool to identify novel proteins that are differentially expressed during mammary development. Proteomics **6**, 5694-5704.

Dy, P., Penzo-Mendez, A., Wang, H., Pedraza, C.E., Macklin, W.B. & Lefebvre, V. (2008) The three SoxC proteins-Sox4, Sox11 and Sox12-exhibit overlapping expression patterns and molecular properties. Nucleic Acids Res. **36**, 3101-3117.

Fornetti, J., Martinson, H.A., Betts, C.B., Lyons, T.R., Jindal, S., Guo, Q.C., Coussens, L.M., Borges, V.F. & Schedin, P. (2014) Mammary gland involution as an

immunotherapeutic target for postpartum breast cancer. J. Mammary Gland Biol. Neoplasia **19**, 213-228.

Gordon, K.E., Binas, B., Chapman, R.S., Kurian, K.M., Clarkson, R.W.E., Clark, A.J., Lane, E.B. & Watson, C.J. (2000) A novel cell culture model for studying differentiation and apoptosis in the mouse mammary gland. Breast Cancer Res. **2**, 222-235.

Hughes, K., Blanck, M., Pensa, S. & Watson, C.J. (2016) Stat3 modulates chloride channel accessory protein expression in normal and neoplastic mammary tissue. Cell Death Dis. **7**, e2398.

Hughes, K. & Watson, C.J. (2018) The multifaceted role of STAT3 in mammary gland involution and breast cancer. Int. J. Mol. Sci. **19**, 1695.

Hughes, K., Wickenden, J.A., Allen, J.E. & Watson, C.J. (2012) Conditional deletion of Stat3 in mammary epithelium impairs the acute phase response and modulates immune cell numbers during post-lactational regression. J. Pathol. **227**, 106-117.

Hunt, S.M.N. & Clarke, C.L. (1999) Expression and hormonal regulation of the Sox4 gene in mouse female reproductive tissues. Biol. Reprod. **61**, 476-481.

Imai, Y., Matsushima, Y., Sugimura, T. & Terada, M. (1991) A simple and rapid method for generating a deletion by PCR. Nucleic Acids Res. **19**, 2785-2785.

Jafarnejad, S.M., Ardekani, G.S., Ghaffari, M. & Li, G. (2013) Pleiotropic function of SRY-related HMG box transcription factor 4 in regulation of tumorigenesis. Cell. Mol. Life Sci. **70**, 2677-2696.

Jena, M.K., Jaswal, S., Kumar, S. & Mohanty, A.K. (2019) Molecular mechanism of mammary gland involution: An update. Dev. Biol. **445**, 145-155.

Jerry, D.J., Dickinson, E.S., Roberts, A.L. & Said, T.K. (2002) Regulation of apoptosis during mammary involution by the p53 tumor suppressor gene. J. Dairy Sci. **85**, 1103-1110.

Jones, F.S. & Jones, P.L. (2000) The tenascin family of ECM glycoproteins: Structure, function, and regulation during embryonic development and tissue remodeling. Dev.

Dyn. **218**, 235-259.

Jones, P.L., Boudreau, N., Myers, C.A., Erickson, H.P. & Bissell, M.J. (1995) Tenascin-C inhibits extracellular matrix-dependent gene expression in mammary epithelial cells. Localization of active regions using recombinant tenascin fragments. J. Cell Sci. **108**, 519-527.

Kondoh, H. & Kamachi, Y. (2010) SOX-partner code for cell specification: Regulatory target selection and underlying molecular mechanisms. Int. J. Biochem. Cell Biol. **42**, 391-399.

Kreuzaler, P.A., Staniszewska, A.D., Li, W.J., Omidvar, N., Kedjouar, B., Turkson, J., Poli, V., Flavell, R.A., Clarkson, R.W.E. & Watson, C.J. (2011) Stat3 controls lysosomal-mediated cell death in vivo. Nat. Cell Biol. **13**, 303-309.

Lund, L.R., Romer, J., Thomasset, N., Solberg, H., Pyke, C., Bissell, M.J., Dano, K. & Werb, Z. (1996) Two distinct phases of apoptosis in mammary gland involution: proteinase-independent and -dependent pathways. Development **122**, 181-193.

Lyons, T.R., O'Brien, J., Borges, V.F., Conklin, M.W., Keely, P.J., Eliceiri, K.W., Marusyk, A., Tan, A.C. & Schedin, P. (2011) Postpartum mammary gland involution drives progression of ductal carcinoma in situ through collagen and COX-2. Nat. Med. **17**, 1109-1115.

Macias, H. & Hinck, L. (2012) Mammary gland development. Wiley Interdiscip. Rev. Dev. Biol. **1**, 533-557.

Malki, S., Boizet-Bonhoure, B. & Poulat, F. (2010) Shuttling of SOX proteins. Int. J. Biochem. Cell Biol. **42**, 411-416.

McCready, J., Arendt, L.M., Glover, E., Iyer, V., Briendel, J.L., Lyle, S.R., Naber, S.P., Jay, D.G. & Kuperwasser, C. (2014) Pregnancy-associated breast cancers are driven by differences in adipose stromal cells present during lactation. Breast Cancer Res. **16**, R2.

Miyoshi, M., Okajima, T., Matsuda, T., Fukuda, M.N. & Nadano, D. (2007) Bystin in human cancer cells: intracellular localization and function in ribosome biogenesis.

Biochem. J. **404**, 373-381.

Muto, A., Sugihara, Y., Shibakawa, M., Oshima, K., Matsuda, T. & Nadano, D. (2018) The mRNA-binding protein Serbp1 as an auxiliary protein associated with mammalian cytoplasmic ribosomes. Cell Biochem. Funct. **36**, 312-322.

Nakamura, M., Tomita, A., Nakatani, H., Matsuda, T. & Nadano, D. (2006) Antioxidant and antibacterial genes are upregulated in early involution of the mouse mammary gland: sharp increase of ceruloplasmin and lactoferrin in accumulating breast milk. DNA Cell Biol. **25**, 491-500.

Nakatani, H., Aoki, N., Nakagawa, Y., Jin-No, S., Aoyama, K., Oshima, K., Ohira, S., Sato, C., Nadano, D. & Matsuda, T. (2006) Weaning-induced expression of a milk-fat globule protein, MFG-E8, in mouse mammary glands, as demonstrated by the analyses of its mRNA, protein and phosphatidylserine-binding activity. Biochem. J. **395**, 21-30.

Nakatani, H., Aoki, N., Okajima, T., Nadano, D., Flint, D. & Matsuda, T. (2010) Establishment of a mammary stromal fibroblastic cell line for in vitro studies in mice of mammary adipocyte dfferentiation. Biol. Reprod. **82**, 44-53.

Neuberg, P. & Kichler, A. (2014) Recent developments in nucleic acid delivery with polyethylenimines. Adv. Genet. **88**, 263-288.

Pan, X., Zhao, J., Zhang, W.N., Li, H.Y., Mu, R., Zhou, T., Zhang, H.Y., Gong, W.L., Yu, M., Man, J.H., Zhang, P.J., Li, A.L. & Zhang, X.M. (2009) Induction of SOX4 by DNA damage is critical for p53 stabilization and function. Proc. Natl. Acad. Sci. USA **106**, 3788-3793.

Penzo-Mendez, A.I. (2010) Critical roles for SoxC transcription factors in development and cancer. Int. J. Biochem. Cell Biol. **42**, 425-428.

Philp, J.A.C., Burdon, T.G. & Watson, C.J. (1996) Differential activation of STATs 3 and 5 during mammary gland development. FEBS Lett. **396**, 77-80.

Rudolph, M.C., McManaman, J.L., Hunter, L., Phang, T. & Neville, M.C. (2003) Functional development of the mammary gland: use of expression profiling and

trajectory clustering to reveal changes in gene expression during pregnancy, lactation, and involution. J. Mam. Gland Biol. Neoplasia **8**, 287-307.

Sargeant, T.J., Lloyd-Lewis, B., Resemann, H.K., Ramos-Montoya, A., Skepper, J. & Watson, C.J. (2014) Stat3 controls cell death during mammary gland involution by regulating uptake of milk fat globules and lysosomal membrane permeabilization. Nat. Cell Biol. **16**, 1057-1068.

Scharer, C.D., McCabe, C.D., Ali-Seyed, M., Berger, M.F., Bulyk, M.L. & Moreno, C.S. (2009) Genome-wide promoter analysis of the SOX4 transcriptional network in prostate cancer cells. Cancer Res. **69**, 709-717.

Stein, T., Salomonis, N. & Gusterson, B.A. (2007) Mammary gland involution as a multi-step process. J. Mammary Gland Biol. Neoplasia **12**, 25-35.

Sugihara, Y., Sadohara, E., Yonezawa, K., Kugo, M., Oshima, K., Matsuda, T. & Nadano, D. (2013) Identification and expression of an autosomal paralogue of ribosomal protein S4, X-linked, in mice: Potential involvement of testis-specific ribosomal proteins in translation and spermatogenesis. Gene **521**, 91-99.

Tiwari, N., Tiwari, V.K., Waldmeier, L., Balwierz, P.J., Arnold, P., Pachkov, M., Meyer-Schaller, N., Schubeler, D., van Nimwegen, E. & Christofori, G. (2013) Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. Cancer Cell **23**, 768-783.

van de Wetering, M., Oosterwegel, M., van Norren, K. & Clevers, H. (1993) Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. EMBO J. **12**, 3847-3854.

Vervoort, S.J., van Boxtel, R. & Coffer, P.J. (2013) The role of SRY-related HMG box transcription factor 4 (SOX4) in tumorigenesis and metastasis: friend or foe? Oncogene **32**, 3397-3409.

Wegner, M. (1999) From head to toes: the multiple facets of Sox proteins. Nucleic Acids Res. **27**, 1409-1420.

Wilson, M. & Koopman, P. (2002) Matching SOX: partner proteins and co-factors of

the SOX family of transcriptional regulators. Curr. Opin. Genet. Dev. **12**, 441-446.

Wiseman, B.S. & Werb, Z. (2002) Stromal effects on mammary gland development and breast cancer. Science **296**, 1046-1049.

Figure legends

Figure 1. Mouse Sox4 and the elevation of its mRNA level in mammary gland involution. (A) Schematic representation of mouse Sox4. Ab, the region used as the immunogen for anti-Sox4 antibody; CT, C-terminal region determining intracellular stability and essential for transactivation; DB, DNA-binding domain; TA, transactivation domain. See text for details. (B) RT-PCR of Sox4 mRNA at L10 and In2. Mammary glands were excised from three different mice at L10 and In2. GAPDH mRNA was used as the internal control. (C) Comparison of Sox4 expression in the mammary gland, ovary, and thymus. RT-PCR was performed as in (B). Mammary glands at In1 were collected from two different mice.

Figure 2. Immunoblotting of Sox4 in involution. (A) Reactivity of anti-mouse Sox4 antibody. The expression vector of FLAG-tagged Sox4 (FLAG-Sox4) or control pCMV-Tag2B was introduced into HEK293T cells. These transfectants were immunoblotted with antibodies against mouse Sox4, a FLAG tag, and α-tubulin (loading control). (B) The increase of Sox4 protein at In2. Mammary glands from three different mice at L10 and In2, and HEK293T expressing FLAG-Sox4 were immunoblotted. MFG-E8 was detected as an involution marker. (C) Preabsorption treatment with the immunogenic peptide of anti-Sox4 antibody abolished the immunoreaction. (D) The expression of Sox4 in the mammary gland, ovary, and thymus was examined at the protein level. Mammary glands at In1 were collected from two different mice.

Figure 3. Nuclear Sox4 was observed in the involuting mammary gland. (A) Cryosections of mammary glands at L10 and In2 were stained with anti-Sox4 antibody and TOTO-3 (nuclei). (B) Specificity verification by peptide blocking. Nuclei of the gland at In2 were stained with propidium iodide. Asterisks, the lumen of the acini. Bars,

50 μm. (C) Immunoblotting of the nuclear fractions prepared from mammary glands. Nucleolar fibrillarin, loading control.

Figure 4. Further immunofluorescent analysis of mammary gland Sox4. (A) Staining of the gland at In2 with antibodies against Sox4 and keratins K8/K18 indicated the expression of Sox4 in epithelial cells. Asterisks, the lumen of the acini. Bar, 15 μm. (B) Apoptosis detection. Mammary gland sections were stained with anti-Sox4 antibody and TOTO-3. Apoptotic cells were identified by the TUNEL assay. Bar, 30 μm. Two mice per time point were finally examined to generate these representative images.

Figure 5. Optimization of vector transfection in mammary HC11. HC11 cells were seeded in 35-mm dishes and transfected with the GFP expression vector (pEGFP-C1) by using the PEI-Max reagent. The amounts (per dish) of vector DNA and PEI-Max are indicated. (A) Microscopic observations. Expression of GFP was observed by fluorescence microscopy under the initial transfection conditions (top two images) and the optimized transfection conditions (middle two images). When the DNA/PEI-Max ratio was too high, insoluble DNA-PEI particles were observed soon after their mixing (phase contrast, bottom image), which resulted in low transfection efficiency (panel B, inset). Bars, 100 μm. (B) Expression of GFP was assessed by immunoblotting with anti-GFP antibody. α-Tubulin, loading control. Note that high amounts of PEI-Max induced cell death by its toxicity, resulting in poor cell recovery and the decrease in tubulin band intensity (bottom).

Figure 6. The effect of Sox4 on the expression of tenascin C mRNA in HC11. (A) HC11 cells were transfected, under the optimized conditions, with the vector for the expression of FLAG-tagged, full-length Sox4 (FLAG-Sox4), FLAG-tagged Sox4 without the CT region (Figure 1A) (FLAG-Sox4-dCT), or control pCMV-Tag2B (mock) and subjected to immunoblotting with anti-FLAG antibody. Asterisk indicates

the position of a non-specific band. α-Tubulin, loading control. (B) Expression of tenascin C mRNA was upregulated by FLAG-Sox4 (left), which was compromised by deletion of the CT region (right). The mRNA expression of tenascin C and control GAPDH was evaluated by semi-quantitative RT-PCR. The control values (left, mock; right, FLAG-Sox4) were set as 1.0. The data represent mean ± s.d. ($n = 3$). Asterisks, $P < 0.05$.

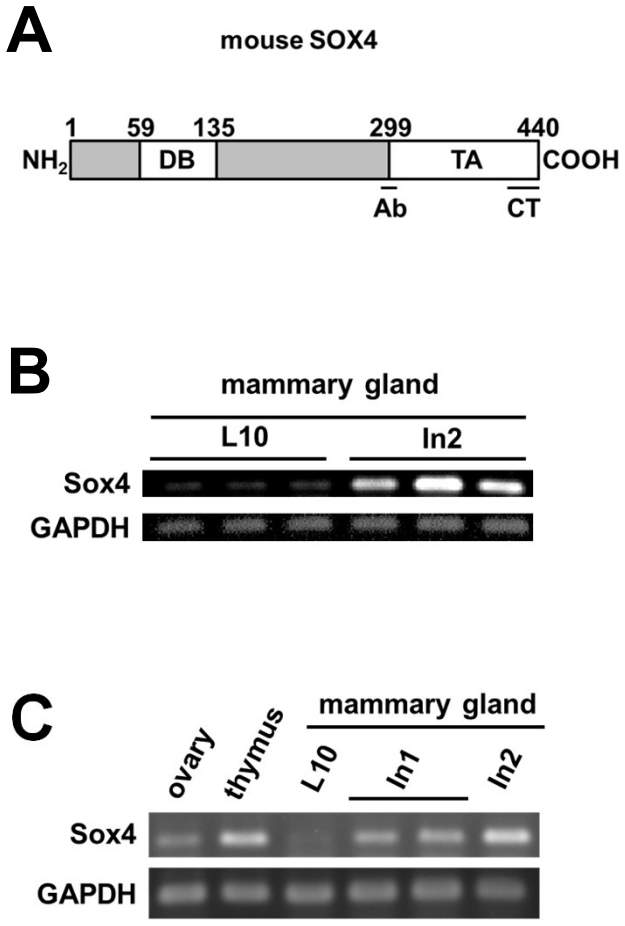

A
mouse SOX4
1
59
135
299
440
NH2
DB
TA
COOH
Ab
CT
B
mammary gland
L10
In2
Sox4
GAPDH
C
mammary gland
ovary
thymus
L10
In1
In2
Sox4
GAPDH

Figure 1

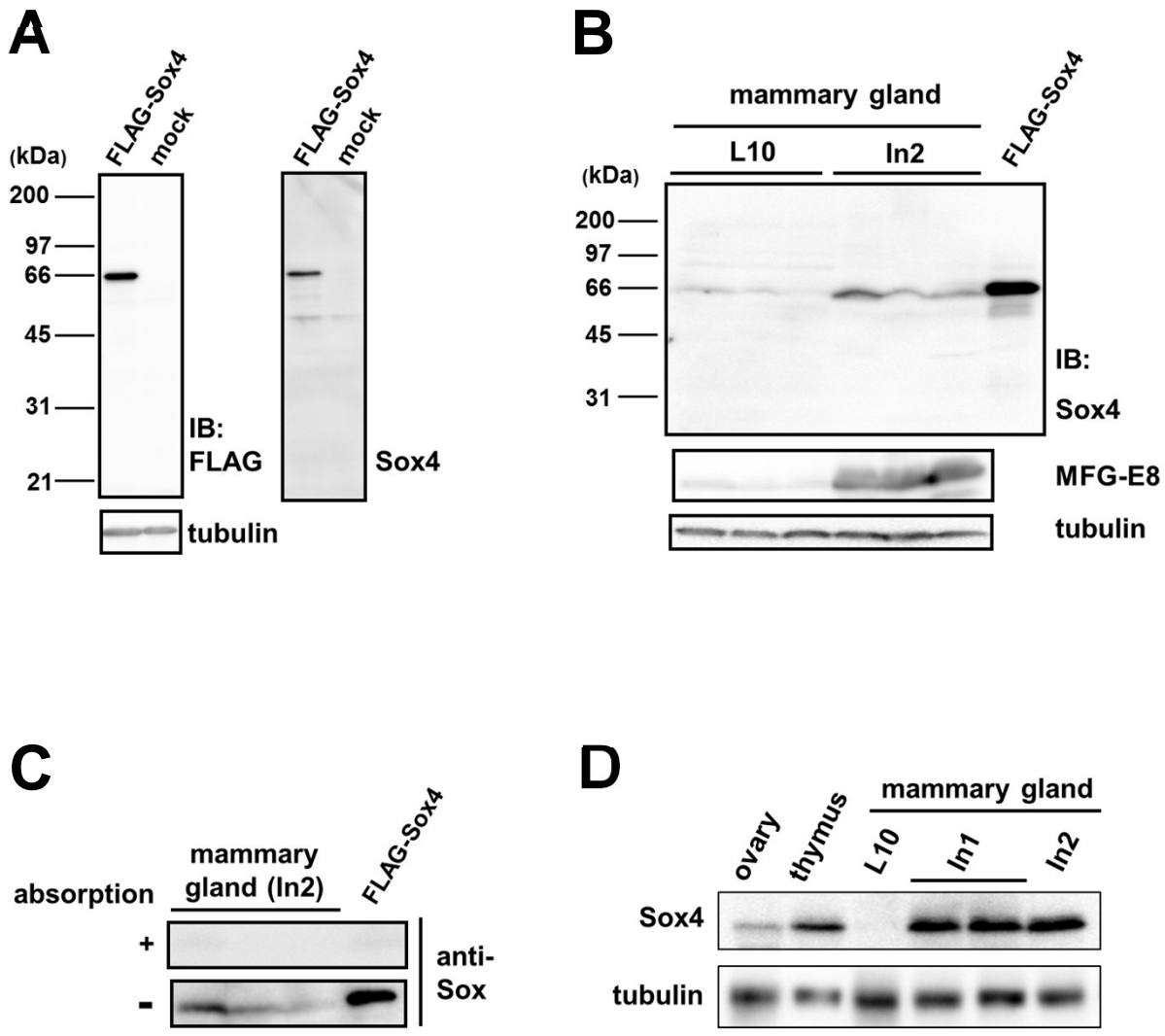

A
FLAG-Sox4
mock
FLAG-Sox4
mock
(kDa)
200
97
66
45
31
21
IB:
FLAG
Sox4
tubulin
B
mammary gland
L10
ln2
FLAG-Sox4
(kDa)
200
97
66
45
31
IB:
Sox4
MFG-E8
tubulin
C
absorption
mammary
gland (ln2)
FLAG-Sox4
+
-
anti-
Sox
D
mammary gland
ovary
thymus
L10
ln1
ln2
Sox4
tubulin

Figure 2

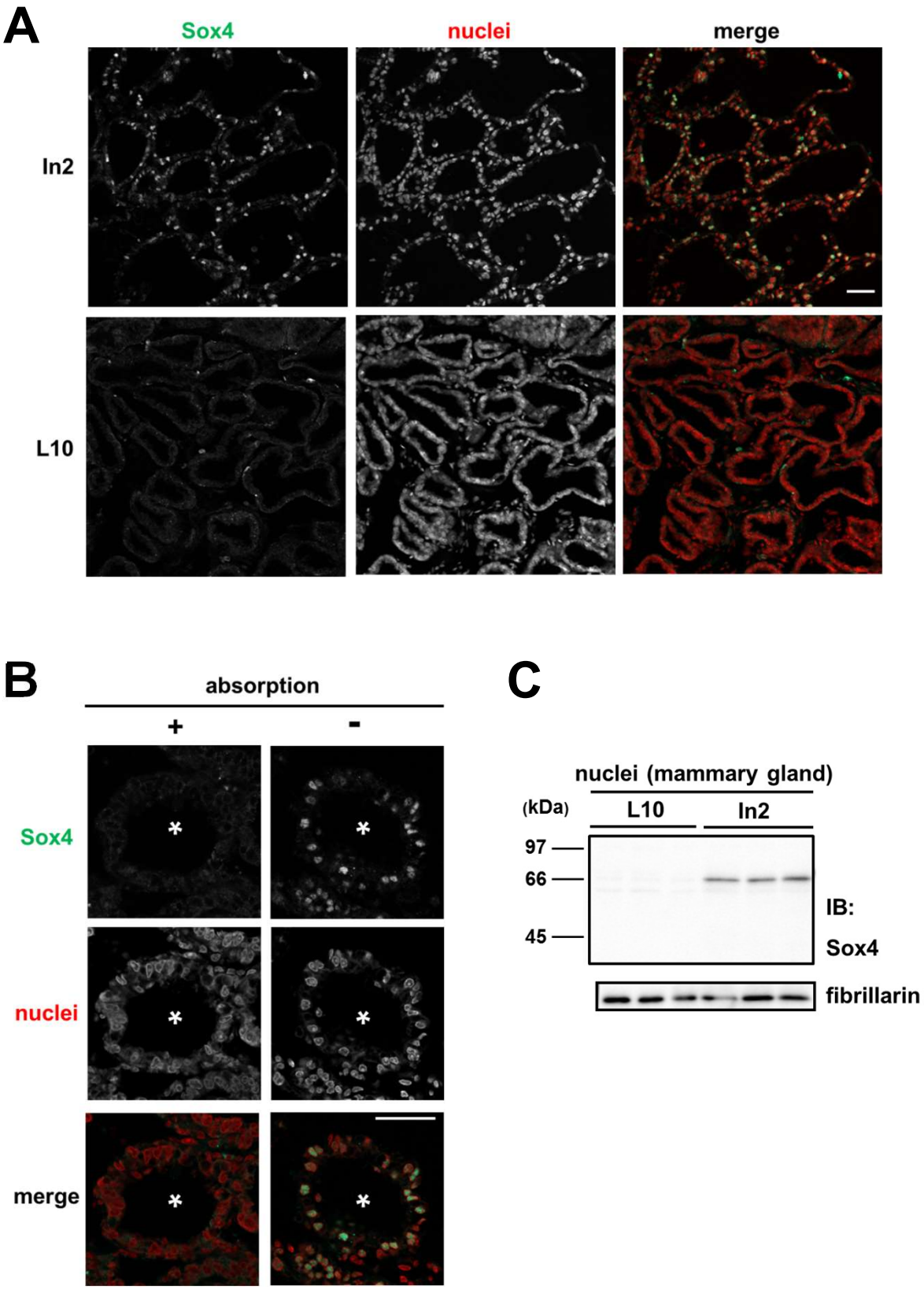

A
Sox4
nuclei
merge
In2
L10
B
absorption
+
-
Sox4
nuclei
merge
C
nuclei (mammary gland)
(kDa)
L10
In2
97
66
45
IB:
Sox4
fibrillarin

Figure 3

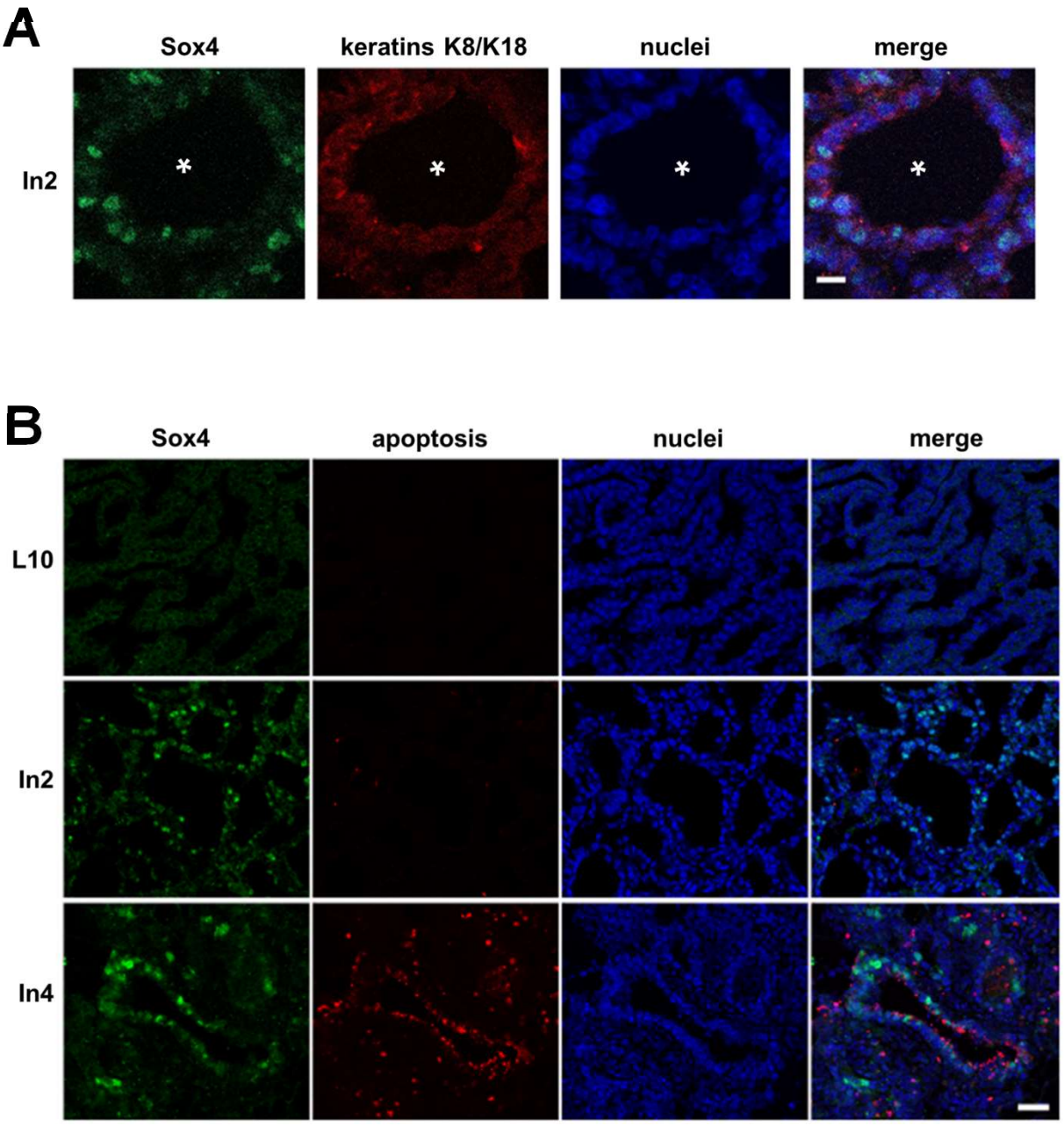

A
Sox4
keratins K8/K18
nuclei
merge
ln2
B
Sox4
apoptosis
nuclei
merge
L10
ln2
ln4

Figure 4

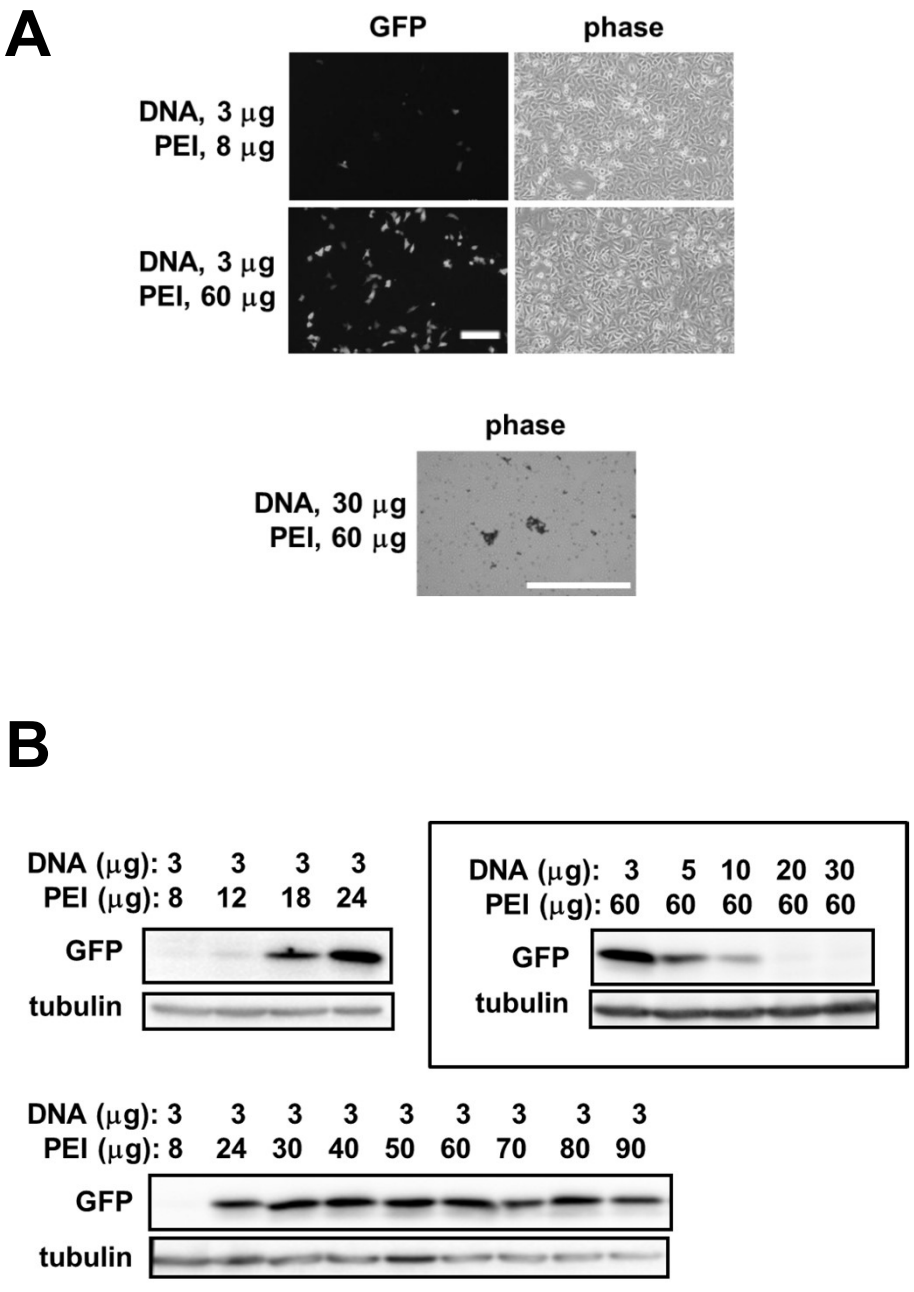

A
GFP
phase
DNA, 3 μg
PEI, 8 μg
DNA, 3 μg
PEI, 60 μg
phase
DNA, 30 μg
PEI, 60 μg
B
DNA (μg): 3 3 3 3
PEI (μg): 8 12 18 24
GFP
tubulin
DNA (μg): 3 5 10 20 30
PEI (μg): 60 60 60 60 60
GFP
tubulin
DNA (μg): 3 3 3 3 3 3 3 3 3
PEI (μg): 8 24 30 40 50 60 70 80 90
GFP
tubulin

Figure 5

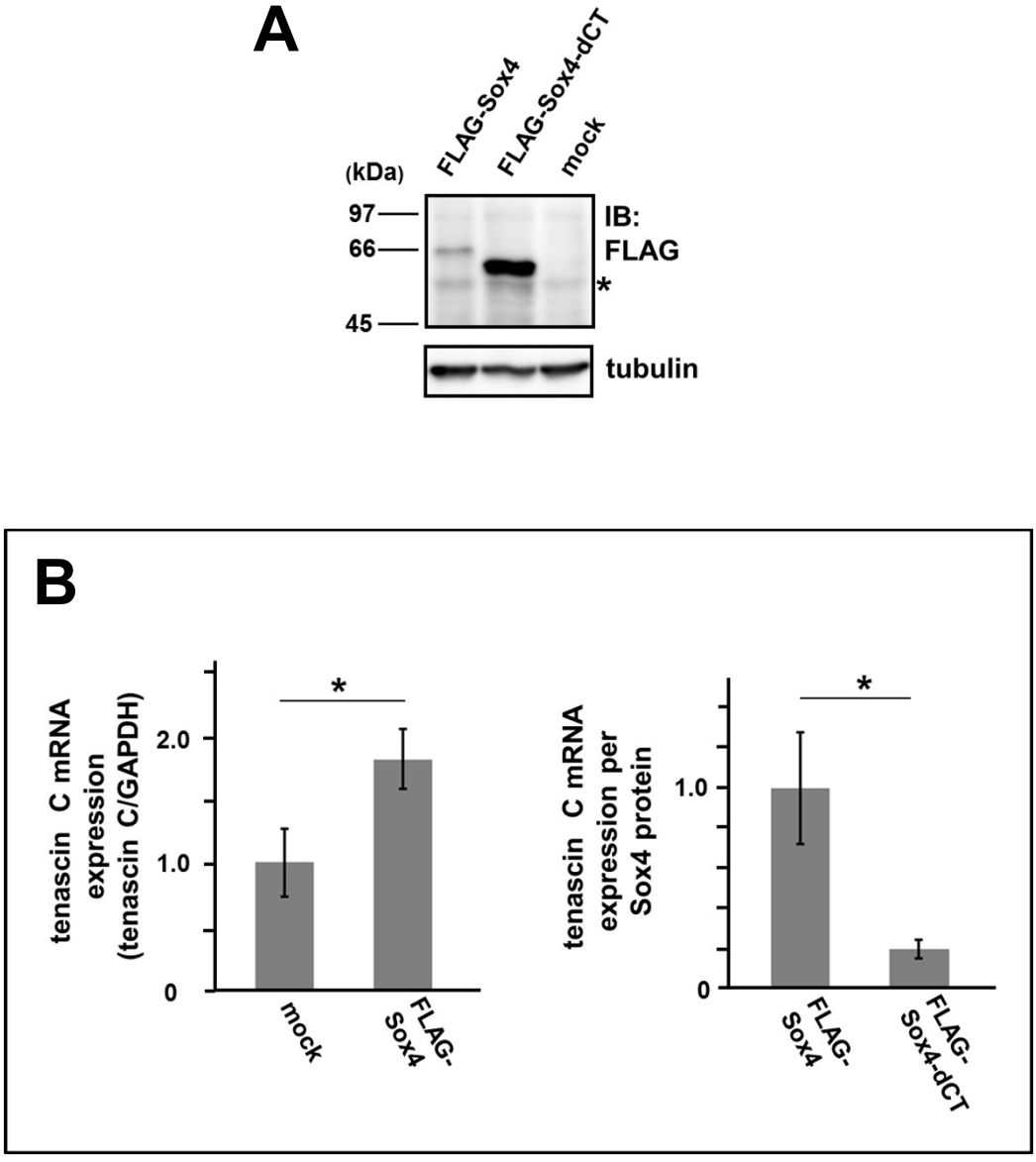
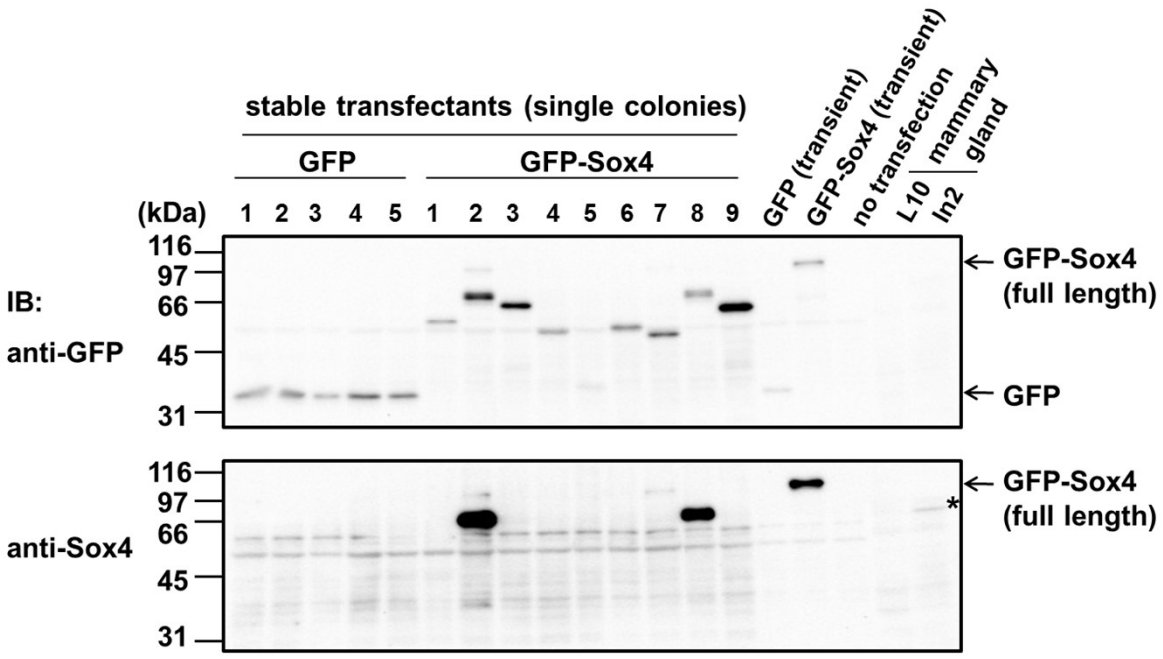

A
FLAG-Sox4
FLAG-Sox4-dCT
mock
(kDa)
97
66
45
IB:
FLAG
*
tubulin
B
tenascin C mRNA expression (tenascin C/GAPDH)
2.0
1.0
0
*
mock
FLAG-Sox4
tenascin C mRNA expression per Sox4 protein
1.0
0
*
FLAG-Sox4
FLAG-Sox4-dCT

Figure 6



Supplementary Figure S1. Unsuccessful establishment of stable transfectants expressing mouse Sox4. MCF-7 cells were transfected with the expression vector of GFP or N-terminally GFP-fused Sox4 (GFP-Sox4) and maintained in the presence of G418. Single colonies exhibiting GFP fluorescence were collected and immunoblotted with antibodies against Sox4 and GFP. Asterisk indicates endogenous Sox4 at In2 in the mouse mammary gland. Note that, while the expression of full-length GFP was detected in all the five GFP transfectants, the incomplete length (most likely C-terminally truncated form) of Sox4 was observed in any of the nine GFP-Sox4 transfectants.