



Title	Isolation and characterization of human breast cancer cells with SOX2 promoter activity
Author(s)	Liang, Shanshan; Furuhashi, Masako; Nakane, Rie; Nakazawa, Seitaro; Goudarzi, Houman; Hamada, Jun-ichi; Iizasa, Hisashi
Citation	Biochemical and biophysical research communications, 437(2), 205-211 https://doi.org/10.1016/j.bbrc.2013.06.038
Issue Date	2013-07-26
Doc URL	http://hdl.handle.net/2115/53246
Type	article (author version)
File Information	Manuscript_20130607.pdf



[Instructions for use](#)

Isolation and characterization of human breast cancer cells with *SOX2* promoter activity

Shanshan Liang, Masako Furuhashi, Rie Nakane¹, Seitaro Nakazawa¹, Houman Goudarzi, Jun-ichi Hamada* and Hisashi Iizasa*

Present address: Division of Stem Cell Biology, Institute for Genetic Medicine, Hokkaido University, Kita 15, Nishi 7, Sapporo 060-0815, Japan.

Running title: *SOX2* promoter plus stem cell

*Corresponding authors at: Division of Stem Cell Biology, Institute for Genetic Medicine, Hokkaido University, Kita 15, Nishi 7, Sapporo 060-0815, Japan. Tel: +81 11 706 6083; Fax: +81 11 706 7870

E-mail address: jhamada@igm.hokudai.ac.jp (J.H.), iizasah@igm.hokudai.ac.jp (H.I.)

¹ These authors contributed to this work equally.

Word count: 4,388

ABSTRACT

Sex determining region Y-box 2 (SOX2) is well known as one of the “stemness” factors and is often expressed in cancers including breast cancer. In this study, we developed a reporter system using fluorescent protein driven by the promoter for *SOX2* gene to detect and isolate living SOX2-positive cells. Using this system, we determined that *SOX2* promoter activities were well correlated with *SOX2* mRNA expression levels in 5 breast cancer cell lines, and that the cell population with positive *SOX2* promoter activity (pSp-T⁺) isolated from one of the 5 cell lines, MCF-7 cells, showed a high *SOX2* protein expression and high sphere-forming activity compared with very low promoter activity (pSp-T^{low/-}). The pSp-T⁺ population expressed higher mRNA levels of several stemness-related genes such as *CD44*, *ABCBI*, *NANOG* and *TWIST1* than the pSp-T^{low/-} population whereas the two populations expressed *CD24* at similar levels. These results suggest that the cell population with *SOX2* promoter activity contains cancer stem cell (CSC)-like cells which show different expression profiles from those of CSC-marker genes previously recognized in human breast cancers.

Keywords: *SOX2* promoter / transcription factor reporter / cancer stem cells / stemness

1. Introduction

The transcription factor sex determining region Y-box 2 (SOX2) is well known to be one of the essential factors for inducing and maintaining the biological nature of stem cells (stemness), namely self-renewal and pluripotency as seen in adult neural progenitor cells (NPCs) [1]. On the other hand, SOX2 is expressed in a variety of cancers such as gastric cancer [2], small cell lung cancer [3], melanoma [4], esophageal squamous cell carcinoma [5] and breast cancer [6], and may be involved in maintaining cancer stem cells (CSCs) which can self-renew, and have pluripotent and tumorigenic ability. However, biological roles of SOX2 in cancer cells remain undetermined. For analyzing the function of SOX2, it is necessary to isolate a small population of living cancer cells expressing SOX2⁺ from a whole cancer cell population; it is extremely difficult to recognize and isolate the live SOX2⁺ cancer cells because SOX2 generally locates in nuclear instead of on cell surface.

Molecular mechanism of *SOX2* gene transcription in cancer is hardly known. It is reported that hypomethylation of *SOX2* promoter enhances *SOX2* gene expression in glioblastoma multiforme [7], and that mRNA expression level of *SOX2* is correlated with *SOX2* hyper-methylation in choriocarcinoma [8]. A study shows that an artificial transcription factor suppressed *SOX2* gene expression by inducing *SOX2* core promoter methylation [9]. These reports suggest that *SOX2* promoter is important for regulating *SOX2* gene transcription in cancers, while the *SOX2* gene enhancers *SRR1* and *SRR2* are important for regulating *SOX2* gene expression in adult NPCs, or in murine embryonic stem cells [10,11].

To isolate SOX2⁺ cells from breast cancer cells, we constructed reporter vectors containing fluorescence protein which was driven by human *SOX2* promoter, and examined the ability of sphere formation of the cells with *SOX2* promoter activity and the expressions of their “stemness”-related genes, using the vectors.

2. Materials and Methods

2.1. Cell culture.

Human breast cancer cell lines MCF-7, T-47D, MDA-MB-231, MDA-MB-436 and ZR-75-1 were obtained from American Type Culture Collection (Manassas, VA). Human teratoma cell line PA-1 was obtained from JCRB cell bank (Ibaraki City, Japan). These cell lines were cultured in Dulbecco's Modified Eagle Medium/ Ham's Nutrient Mixture F-12 (D-MEM/Ham's F-12) (Wako Pure Chemical, Osaka, Japan) supplemented with heat-inactivated 10% fetal bovine serum (Cambrex, East Rutherford, NJ), 100 units/ml penicillin and 100 µg/ml streptomycin sulfate (both from Invitrogen, Carlsbad, CA), and maintained at 37°C, 5% CO₂.

2.2. Construction of SOX2 promoter-driven fluorescent protein gene and luciferase gene reporter vectors.

Human *SOX2* promoter region (-789 to +253) and two regulatory regions (*SRR1* and *SRR2*) (NT_005612.16) were amplified from genomic DNA of MCF-7 cell line by using specific primers (*SOX2* promoter FW: 5'-GGTACCGGCCAAAGAGCTGAGTTGGA-3', RV: 5'-AAGCTTGAGGCA AACTGGAATCAGGATC-3', *SRR1* FW: 5'-TTTTGGAGCCTACAGTTGAC-3', RV: 5'-GGTGAGA ACTAGCCAAGCATCT-3', *SRR2* FW: 5'-CCAGTTCCCAATTCTAAGCAA-3', RV: 5'-CCAGGAT GATCATTTGGACT-3'), and cloned into pGEM-T easy vector (Promega, Madison, WI). *SOX2* promoter and/or *SRRs* were inserted into a luciferase reporter vector (pGL3-Basic; Promega) and the constructed vectors were designated as pSp-L, pSI-Sp-L, pSp-L-SII and pSI-Sp-L-SII reporter vector. For detecting the live cells expressing *SOX2*, firefly luciferase gene was replaced to a gene encoding tdTomato fluorescent protein which was cloned from ptdTomato-C1 plasmid vector (Clontech Laboratories, Mountain View, CA) by PCR using the primers (FW: 5'-CCATGGTGAGCAAGGGCGA G-3', RV: 5'-TCTAGAATTACTTGTACAGCTCGTCCATGC-3') (pSp-T reporter vector). When the cells reached 70% confluence, they were re-suspended in Opti-MEM medium (Invitrogen, Carlsbad, CA) and pSp-T reporter vector or control vector ptdTomato-C1 was transfected into the cells by electroporator (CUY21SC, Nepa Gene, Ichikawa, Japan). And each luciferase reporter vector and internal control pRL-TK vector (Promega) were co-transfected into the cells. Twenty four hours after transfection, luciferase activities were measured by using the Dual-Luciferase® Reporter Assay System (Promega) with a luminometer MiniLumat (Berthold, Bad Wildbad, Germany). Relative luciferase activity was expressed as the firefly luciferase activity normalized by the Renilla luciferase activity.

2.3. Flow cytometry and cell sorting.

SOX2 promoter activity was monitored and the cells with *SOX2* promoter activity were sorted out with flow cytometer FACS Aria II (BD Biosciences, San Jose, CA). The cells with tdTomato signal were sorted out, seeded in a plate for culture. Preparation for flow cytometric analysis of cells expressing *SOX2* was carried out as follows. The cells transfected with pSp-T reporter vector were fixed with 4% formaldehyde for 10 minutes at 37°C and permeabilized by ice-cold 100% methanol for 30 minutes. Then the cells were treated with 3% bovine serum albumin for 1 hour and incubated with rabbit anti-human *SOX2* monoclonal antibody (#3579; Cell signaling Technology, Danvers, MA) for 1 hour at room temperature. After washing with PBS, the cells were incubated with CF488A-conjugated goat anti-rabbit IgG antibody (Biotium, Hayward, CA) for 1 hour at room temperature.

2.4. Western blotting.

The cells were lysed with RIPA buffer [25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS containing proteinase inhibitor (Complete Proteinase inhibitor Cocktail, Roche Diagnostics, Mannheim, Germany)]. Protein extracts were fractionated on 14% polyacrylamide gels and transferred onto PVDF membranes (Immobilon P: Millipore, Billerica, MA). Rabbit monoclonal antibody to human SOX2 and mouse monoclonal antibody to human actin (MAB1501: Millipore) were used as primary antibodies, and horseradish peroxidase (HRP)-Linked sheep F (ab') 2 Fragment of antibody to mouse IgG and HRP-Linked donkey F (ab') 2 Fragment of antibody to rabbit IgG (GE Healthcare Bioscience, Pittsburgh, PA) were used as secondary antibodies. Each protein band was detected by using Western Chemiluminescent HRP Substrate (Millipore) and visualized by LAS1000mini (Fujifilm, Tokyo, Japan).

2.5. Sphere formation.

The cells were seeded (100 cells/well) on ultra-low-attachment 96-well plate (Sumitomo Bakelite, Tokyo, Japan) by using FACS Aria II and cultured in a sphere-forming medium (D-MEM/Ham's F-12, 20 ng/ml of epidermal growth factor, 20 ng/ml of basic fibroblast growth factor, 100 units/ml penicillin and 100 µg/ml of streptomycin sulfate) for 7–10 days. Sphere-forming ability was evaluated from the percentage of the wells where a sphere was formed against the wells where the cells had been seeded [mean ± standard deviation (SD), n = 3 plates].

2.6. Quantitative real-time reverse transcription-polymerase chain reaction (qRT-PCR)

qRT-PCR was performed with an ABI Prism 7900HT (Applied Biosystems, Grand Island, NY) using 2x Quantifast SYBR Green PCR Master mix (QIAGEN, Valencia, CA) with 1 µM primers. Specific primer sets used were human *SOX2* (FW: 5'-TGGACAGTTACGCGCACAT-3', RV: 5'-CGAGTAGG ACATGCTGTAGGT-3') and *ACTB* (β-actin) (FW: 5'-TTGCCGACAGGATGCAGAA-3' and RV: 5'-GCCGATCCACACGGAGTACT-3'). Each target and a standard *ACTB* gene were analyzed in three independent qRT-PCR assays. Thermal cycling was initiated with an initial activation step for 5 min at 95°C, followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. Relative expression of the target gene was analyzed by the $\Delta\Delta C_t$ method. The ratios of mRNA levels were expressed relative to those of the control group.

2.6. RNA extraction and RT-PCR.

Total RNA was extracted from the cultured cells by using TRIZOL reagent (Invitrogen) and treated with DNase I (Invitrogen). First-strand cDNA was synthesized with 1 µg of total RNA and oligo dT (15) primer or *SOX2* sense specific primer (5'-GTTGCCTGGCTTCTCTTTG -3') or *SOX2* antisense specific primer (5'-GCTGATTGGTCGCTAGAAAC-3') by using Superscript III according to the manufacturer's instructions (Invitrogen). The synthesized cDNA was then amplified by PCR using Go-Taq® DNA Polymerase system (Promega) and specific primers for *CD44* (FW: 5'-AGACGAAGACAG TCCCTGGATCAC-3', RV: 5'-TGTGTTTGC TCCACCTTCTTGACTC-3'), *CD24* (FW: 5'-CTCCTA CCCACGCAGATTTATTC-3', RV: 5'-TGG TGGCATTAGTTGGATTTGG-3'), *ABCG2* (FW: 5'-TGGCACTGGCCATAGCAGCA-3', RV: 5'-AGCCATGACAGCCAAGATGCAA-3'), *ABCB1* (FW: 5'-AGGTGCTGCTTTTCTGCTGA-3', RV: 5'-GCCTGTCCAACACTAAAAGCCC-3'), *ABCC1* (FW: 5'-GAGGCGCCCTGGCAAATCCA-3', RV: 5'-GCCCCACGATGCCGACCTTT-3'), *NANOG* (FW:

5'-AAGGTCCCGGTCAAGAAACAG-3', RV: 5'-CTTCTGCGTCACACCATTGC-3'), *POU5F1* (FW: 5'-GGTGGAGAGCAACTCCCGATG-3', RV: 5'-CCAGGGTGATCCTCTTCTGC-3'), *KLF4* (FW: 5'-GGGAGAAGACACTGCGTCAA-3', RV: 5'-TCCAGGTCCAGGAGATCGTT-3'), *MYCBP* (FW: 5'-CGGATTCTCTGCTCTCCTCG-3', RV: 5'-CCTCCAGCAGAAGGTGATCC-3'), *CDH1* (FW: 5'-AACGCCCCCATAACCAGAACC-3', RV: 5'-AACAGCAAGAGCAGCAGAATCAGA-3'), *CDH2* (FW: 5'-GAATGGATGAAAGACCCATCC-3', RV: 5'-AGCTTCTGAATGCTTTTTGGG-3'), *ZEB1* (FW: 5'-CAGCCCCTGCAGTCCAAGAAC-3', RV: 5'-CCTGCCTCTTCTGCTCTGT-3'), *ZEB2* (FW: 5'-AGCAAAGGAGAAAGTACCAGCG-3', RV: 5'-GAGTGAAGCCTTGAGTGCTCG-3'), *TWIST1* (FW: 5'-AGCTGAGCAAGATTCAGACCCTC-3', RV: 5'-CCCGTCTGGGAATCACTGTC-3'), *VIM* (FW: 5'-GCAGGCTCAGATTCAGGGAAC-3', RV: 5'-GCTTCAACGGCAAAGTTCTC-3') and *GAPDH* (FW: 5'-TGAAGGTCGGAGTCAACGGATTTGGT-3', RV: 5'-CATGTGGGCCATGAGGTCCACCAC-3').

2.7. Statistical Analysis.

Experiments were run in triplicate or quadruplicate and repeated in a minimum of three independent trials. Data are presented as means \pm SD. Two-tailed Student's t-tests were conducted where the minimum level of significance was $p < 0.05$. Correlation coefficient (R^2) was calculated by Pearson product moment correlation coefficient.

3. Results

3.1. Correlation between *SOX2* promoter activities and endogenous *SOX2* mRNA expression levels in the human breast cancer cell lines

The promoter region -789 to +253 in human *SOX2* gene is highly conserved in human, cow, wolf, rat and mouse [12-14]. *SOX2* promoter activities in MCF-7, T-47D, MDA-MB-231, MDA-MB-436 and ZR-75-1 cells were 87.9, 39.6, 2.3, 7.9 and 13.5, respectively (Fig. 1A). *SOX2* mRNA expression levels in these breast cancer cell lines were detected by qRT-PCR (Fig. 1B); they correlated with the *SOX2* promoter activities in these cell lines (Fig. 1C, $R^2 = 0.97894$). On the other hand, *SOX2* gene enhancers SRRs did not increase pSp-L reporter activities in MCF-7, T-47D or MDA-MB-231 (Fig. 1D).

3.2. Correlation between expressions of *SOX2* proteins and *SOX2* promoter activities in the breast cancer cell lines.

Next we used pSp-T vector which had a reporter gene encoding tdTomato fluorescent protein, in order to detect the live cells with *SOX2* promoter activity. Among MCF-7 cells which were transfected with pSp-T vector, a small number of tdTomato-positive (pSp-T⁺) cells were observed whereas almost all the cells expressed tdTomato fluorescence when transfected with control vector ptdTomato-C1 (Fig. 2A). We examined the ratios of pSp-T⁺ cells in three kinds of breast cancer cell lines by FACS analysis. It revealed that MCF-7, T-47D and MDA-MB-231 cell lines contained 9.2%, 4.5% and 3.2% pSp-T⁺ cells, respectively (Fig. 2B).

To examine *SOX2* protein expression levels in pSp-T⁺ cells, we established two cloned pSp-T⁺ sublines from MCF-7 cells, and designated them as Clone 1 and Clone 2. After the two sublines were cultured for two months, they were fixed and subjected to immunofluorescence staining with anti-*SOX2*

antibody and CF488A-conjugated secondary antibody; their fluorescence intensity was analysed by flow cytometry. It revealed that Clone 1 and Clone 2 sublines contained 62% and 50% pSp-T⁺ cells, respectively. Clone 1 contained 46% pSp-T⁺/SOX2 positive (SOX2⁺) cells, 16% pSp-T⁺/SOX2 negative (SOX2⁻) cells, 16% pSp-T⁻/SOX2⁺ cells and 22% pSp-T⁻/SOX2⁻ cells whereas Clone 2 contained 32% pSp-T⁺/SOX2⁺ cells, 18% pSp-T⁺/SOX2⁻ cells, 10% pSp-T⁻/SOX2⁺ cells and 40% pSp-T⁻/SOX2⁻ cells (Fig. 2C and supplemental Fig. 2).

We further compared the expression levels of SOX2 protein and promoter activities among the cell populations with different tdTomato fluorescence intensities by Western blotting and luciferase reporter assay. We transfected pSp-L reporter vector into Clone 1-derived cells with different levels of tdTomato intensities and examined SOX2 protein expression in pSp-T⁺ cells by Western blotting (Fig. 2D). SOX2 protein expression levels were higher in pSp-T^{medium} and ^{high} cells than in pSp-T^{low/-} cells (Fig. 2D). SOX2 promoter activities were consistent with SOX2 protein levels in each isolated cell population (Fig. 2E).

3.3. Sphere-forming ability of MCF-7 cells with high SOX2 promoter activity

We isolated pSp-T⁺ and pSp-T^{low/-} cells from Clone-1 and -2 cells and examined their sphere-forming ability, one of the phenotypic characters of cancer stem cells. pSp-T⁺ cells were highly sphere-forming compared with parent, mock or pSp-T⁻ cells (Fig. 3A and B).

3.4. mRNA expression profile of stemness-related genes in MCF-7 cells with high SOX2 promoter activity

Several reports indicate that CD44⁺/CD24^{low} cell population has high CSC activity in human breast cancer tissues and cell lines [15]. CD44 mRNA expression level of pSp-T⁺ cells was higher than that of pSp-T⁻ cells whereas CD24 mRNA expression levels were not different between them (Fig. 4A). It is known that side population sorted with FACS on the basis of high efflux of DNA-binding dye Hoechst 33342 is rich in CSCs in some types of cancers [16]. As members of the ATP-binding cassette family such as ABCG2, ABCB1 and ABCC1 are involved in the efflux of Hoechst 33342 [16]. In our experiment, expression of ABCB1 was high in pSp-T⁺ cells, but those of ABCG2 and ABCC1 did not differ between pSp-T⁺ cells and pSp-T^{low/-} cells. Many transcription factors such as Octamer-binding transcription factor 4 (OCT4), Kruppel-like factor 4 (KLF4), NANOG and C-MYC are known to be associated with induction or maintenance of stemness of CSCs [17-19]. In fact, pSp-T⁺ cells highly expressed NANOG, C-MYC compared with pSp-T^{low/-} cells or parent cells (Fig. 4C). SOX2 and KLF4 expression levels in pSp-T⁺ cells were lower than those in pSp-T^{low/-} cells although SOX2 protein expression level of pSp-T⁺ cells was higher than that of pSp-T^{low/-} cells (Fig. 3A)

There is a possibility that PCR products from the cDNA templates reverse-transcribed with oligo dT primers contain both sense and antisense mRNA-derived products. Therefore, to distinguish the sense mRNA-derived PCR products from the antisense mRNA-derived ones, we performed RT-PCR with RT primers specific to sense or antisense SOX2 mRNA. As shown in Fig. 4E, antisense SOX2 mRNA was expressed in pSp-T⁺ cells but not in pSp-T^{low/-} cells. Moreover, the expression levels of EMT-related genes such as TWIST1 and VIM (vimentin) were increased in pSp-T⁺ cells compared with pSp-T^{low/-} or parent (Fig. 4D).

4. Discussion

SOX2, one of the transcription factors pertaining to stemness, is highly expressed in a variety of cancers such as gastric, esophageal, breast, or small cell lung cancer and melanoma [2-6]. However, the function of SOX2 in cancer cells remains unclear since it is difficult to isolate living cells expressing SOX2 from cancer tissues or cell lines. And the reason why it is difficult is that no specific antibody to SOX2 can reach the protein located in the nuclei of the live cells. In this study, we utilized *SOX2* promoter-driven reporter as a strategy to detect SOX2-expressing cells, and found that *SOX2* promoter activities well correlated with expression levels of SOX2 mRNA in five human breast cancer cell lines (Fig. 1). Wu et al. reported that there was no difference in SOX2 expression levels between MCF-7 cells with high enhancer activity and those with no activity, from their reporter assay using GFP reporter driven by *SRR2*, a *SOX2* enhancer [20]. Our reporter analyses revealed that *SRRs* was essential for transactivation of luciferase reporter gene in a teratoma cell line but not in breast cancer cell lines (Fig. 1C and supplemental Fig. 1). These findings indicated that *SOX2* transcription of human breast cancer cells was up-regulated mainly in a promoter- but not *SRRs*-dependent manner.

We next isolated two cloned sublines cells (Clone 1 and 2) with *SOX2* promoter activity from MCF-7 cell line by using a fluorescence protein reporter gene. After subculture of the cloned sublines for two months, each subline showed decreased ratios of fluorescence-positive cells to whole cell population, suggesting that the cells with *SOX2* promoter activity generated cells without the activity. We further sorted out the cell populations with low or no promoter activity and the one with positive *SOX2* promoter activity from the cloned sublines cultured for two months, and analyzed their stemness-related properties. Although the cell population with low or no *SOX2* promoter activity showed low SOX2 expression at protein level compared with the other population, the two cell populations showed similar expression levels of *SOX2* mRNA. This discrepancy may be explained by different levels in antisense mRNA production of *SOX2* between the population with high SOX2 activity and the one with low or no activity if we infer from the report that antisense mRNA functions as an activator of translation of some genes.

The cells with positive *SOX2* promoter activity showed higher sphere-forming activity, a typical phenotype of CSCs, than those with low or no *SOX2* promoter activity or parent cells. Several molecular markers of CSC-like cells are known in human breast cancer, such as *CD44*, *ABCB1*, *ABCG2*, *NANOG*, and *TWIST1*. RT-PCR analysis of these genes presented that the cells with high SOX2 promoter activity had high expressions of *CD44*, *NANOG*, *C-MYC*, *ABCB1* and *TWIST1* compared with those with its low or no activity. The expression level of *CD24* of the cells with high *SOX2* promoter activity was similar to that of the cells with low activity. These findings suggest that the cell population with high SOX2 promoter activity consists of a high concentration of CSC-like cells which are different from those already recognized in CSC population of breast cancer.

Acknowledgements

The authors thank Ms. Masako Yanome for her help in preparing the manuscript. This study was supported by a grant from a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (Grant Number #22659079, # 23590451 and #23590443). The authors have no financial relationship to disclose.

References

- [1] A. Sarkar, K. Hochedlinger, The sox family of transcription factors: versatile regulators of stem and progenitor cell fate, *Cell Stem Cell* 12 (2013) 15-30.
- [2] X.L. Li, Y. Eishi, Y.Q. Bai, H. Sakai, Y. Akiyama, M. Tani, T. Takizawa, M. Koike, Y. Yuasa, Expression of the SRY-related HMG box protein SOX2 in human gastric carcinoma, *Int. J. Oncol.* 24 (2004) 257-263.
- [3] C.M. Rudin, S. Durinck, E.W. Stawiski, J.T. Poirier, Z. Modrusan, D.S. Shames, E.A. Bergbower, Y. Guan, J. Shin, J. Guillory, C.S. Rivers, C.K. Foo, D. Bhatt, J. Stinson, F. Gnad, P.M. Haverty, R. Gentleman, S. Chaudhuri, V. Janakiraman, B.S. Jaiswal, C. Parikh, W. Yuan, Z. Zhang, H. Koeppen, T.D. Wu, H.M. Stern, R.L. Yauch, K.E. Huffman, D.D. Paskulin, P.B. Illei, M. Varellagarcia, A.F. Gazdar, F.J. de Sauvage, R. Bourgon, J.D. Minna, M.V. Brock, S. Seshagiri, Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer, *Nat. Genet.* 44 (2012) 1111-1116.
- [4] S.D. Girouard, A.C. Laga, M.C. Mihm, R.A. Scolyer, J.F. Thompson, Q. Zhan, H.R. Widlund, C.W. Lee, G.F. Murphy, SOX2 contributes to melanoma cell invasion, *Lab. Invest.* 92 (2012) 362-370.
- [5] A.J. Bass, H. Watanabe, C.H. Mermel, S. Yu, S. Perner, R.G. Verhaak, S.Y. Kim, L. Wardwell, P. Tamayo, I. Gat-Viks, A.H. Ramos, M.S. Woo, B.A. Weir, G. Getz, R. Beroukhi, M. O'Kelly, A. Dutt, O. Rozenblatt-Rosen, P. Dziunycz, J. Komisarof, L.R. Chirieac, C.J. Lafargue, V. Scheble, T. Wilbertz, C. Ma, S. Rao, H. Nakagawa, D.B. Stairs, L. Lin, T.J. Giordano, P. Wagner, J.D. Minna, A.F. Gazdar, C.Q. Zhu, M.S. Brose, I. Ceconello, U.R. Jr, S.K. Marie, O. Dahl, R.A. Shivdasani, M.S. Tsao, M.A. Rubin, K.K. Wong, A. Regev, W.C. Hahn, D.G. Beer, A.K. Rustgi, M. Meyerson, SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas, *Nat. Genet.* 41 (2009) 1238-1242.
- [6] S.M. Rodriguez-Pinilla, D. Sarrio, G. Moreno-Bueno, Y. Rodriguez-Gil, M.A. Martinez, L. Hernandez, D. Hardisson, J.S. Reis-Filho, J. Palacios, Sox2: a possible driver of the basal-like phenotype in sporadic breast cancer, *Mod. Pathol.* 20 (2007) 474-481.
- [7] M.M. Alonso, R. Diez-Valle, L. Manterola, A. Rubio, D. Liu, N. Cortes-Santiago, L. Urquiza, P. Jauregi, A. Lopez de Munain, N. Sampron, A. Aramburu, S. Tejada-Solis, C. Vicente, M.D. Odero, E. Bandres, J. Garcia-Foncillas, M.A. Idoate, F.F. Lang, J. Fueyo, C. Gomez-Manzano, Genetic and epigenetic modifications of Sox2 contribute to the invasive phenotype of malignant gliomas, *PLoS One* 6 (2011) e26740.
- [8] A.S. Li, M.K. Siu, H. Zhang, E.S. Wong, K.Y. Chan, H.Y. Ngan, A.N. Cheung, Hypermethylation of SOX2 gene in hydatidiform mole and choriocarcinoma, *Reprod. Sci.* 15 (2008) 735-744.
- [9] S. Stolzenburg, M.G. Rots, A.S. Beltran, A.G. Rivenbark, X. Yuan, H. Qian, B.D. Strahl, P. Blancafort, Targeted silencing of the oncogenic transcription factor SOX2 in breast cancer, *Nucleic Acids Res.* 40 (2012) 6725-6740.
- [10] M. Tomioka, M. Nishimoto, S. Miyagi, T. Katayanagi, N. Fukui, H. Niwa, M. Muramatsu, A. Okuda, Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex, *Nucleic Acids Res.* 30 (2002) 3202-3213.
- [11] M. Sikorska, J.K. Sandhu, P. Deb-Rinker, A. Jezierski, J. Leblanc, C. Charlebois, M. Ribocco-Lutkiewicz, M. Bani-Yaghoob, P.R. Walker, Epigenetic modifications of SOX2 enhancers, SRR1 and SRR2, correlate with in vitro neural differentiation, *J. Neurosci. Res.* 86 (2008) 1680-1693.
- [12] K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, *Nucleic Acids Res.* 30 (2002) 3059-3066.
- [13] K. Katoh, K. Kuma, H. Toh, T. Miyata, MAFFT version 5: improvement in accuracy of multiple sequence alignment, *Nucleic Acids Res.* 33 (2005) 511-518.
- [14] Y. Katoh, M. Katoh, Comparative genomics on SOX2 orthologs, *Oncol. Rep.* 14 (2005) 797-800.

- [15] D. Ponti, A. Costa, N. Zaffaroni, G. Pratesi, G. Petrangolini, D. Coradini, S. Pilotti, M.A. Pierotti, M.G. Daidone, Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties, *Cancer Res.* 65 (2005) 5506-5511.
- [16] L. Patrawala, T. Calhoun, R. Schneider-Broussard, J. Zhou, K. Claypool, D.G. Tang, Side population is enriched in tumorigenic, stem-like cancer cells, whereas ABCG2⁺ and ABCG2⁻ cancer cells are similarly tumorigenic, *Cancer Res.* 65 (2005) 6207-6219.
- [17] J.W. Jung, S.B. Park, S.J. Lee, M.S. Seo, J.E. Trosko, K.S. Kang, Metformin represses self-renewal of the human breast carcinoma stem cells via inhibition of estrogen receptor-mediated OCT4 expression, *PLoS One* 6 (2011) e28068.
- [18] F. Yu, J. Li, H. Chen, J. Fu, S. Ray, S. Huang, H. Zheng, W. Ai, Kruppel-like factor 4 (KLF4) is required for maintenance of breast cancer stem cells and for cell migration and invasion, *Oncogene* 30 (2011) 2161-2172.
- [19] Y. Chen, L. Shi, L. Zhang, R. Li, J. Liang, W. Yu, L. Sun, X. Yang, Y. Wang, Y. Zhang, Y. Shang, The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer, *J. Biol. Chem.* 283 (2008) 17969-17978.
- [20] F. Wu, J. Zhang, P. Wang, X. Ye, K. Jung, K.M. Bone, J.D. Pearson, R.J. Ingham, T.P. McMullen, Y. Ma, R. Lai, Identification of two novel phenotypically distinct breast cancer cell subsets based on Sox2 transcription activity, *Cell Signal.* 24 (2012) 1989-1998.

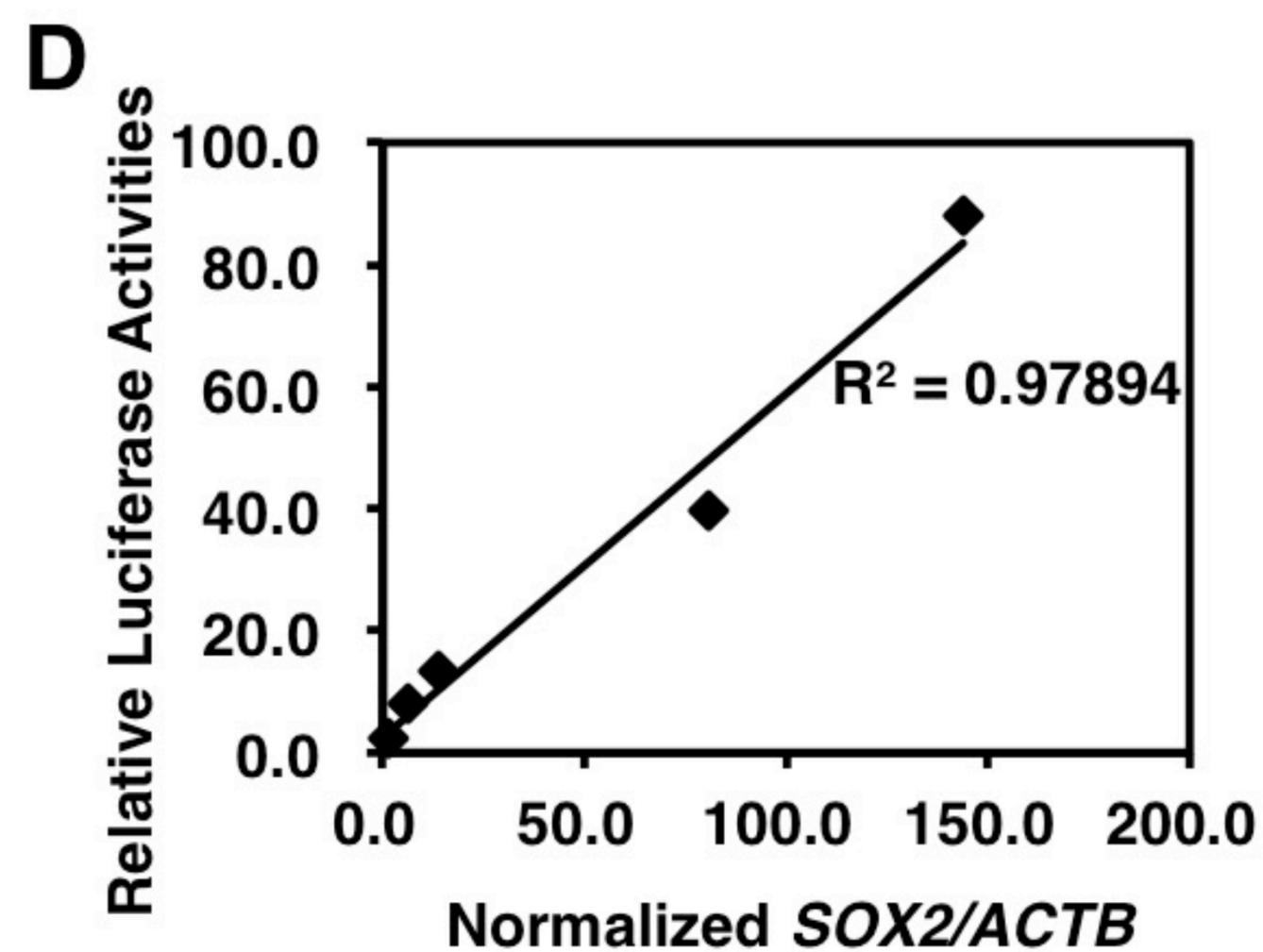
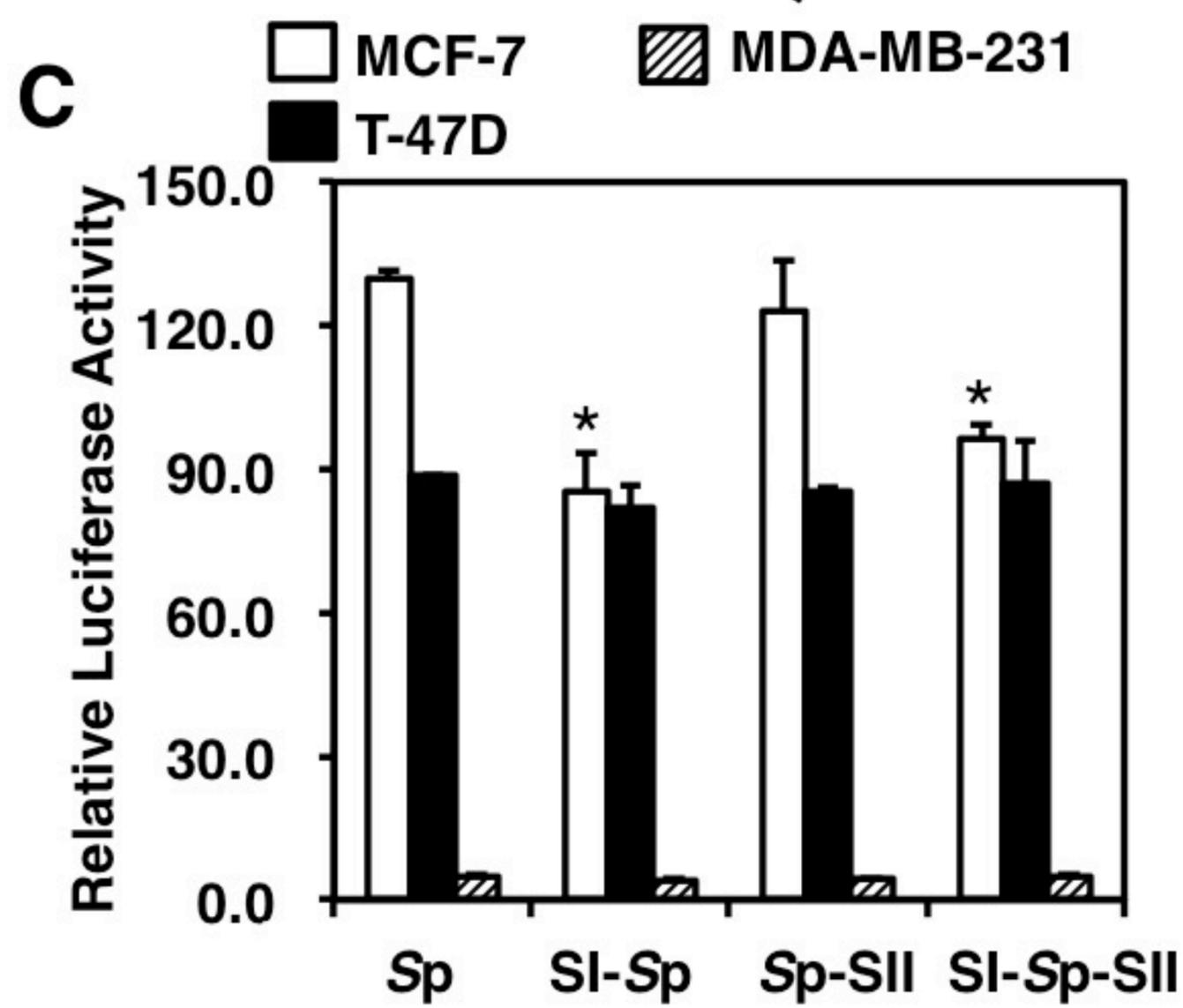
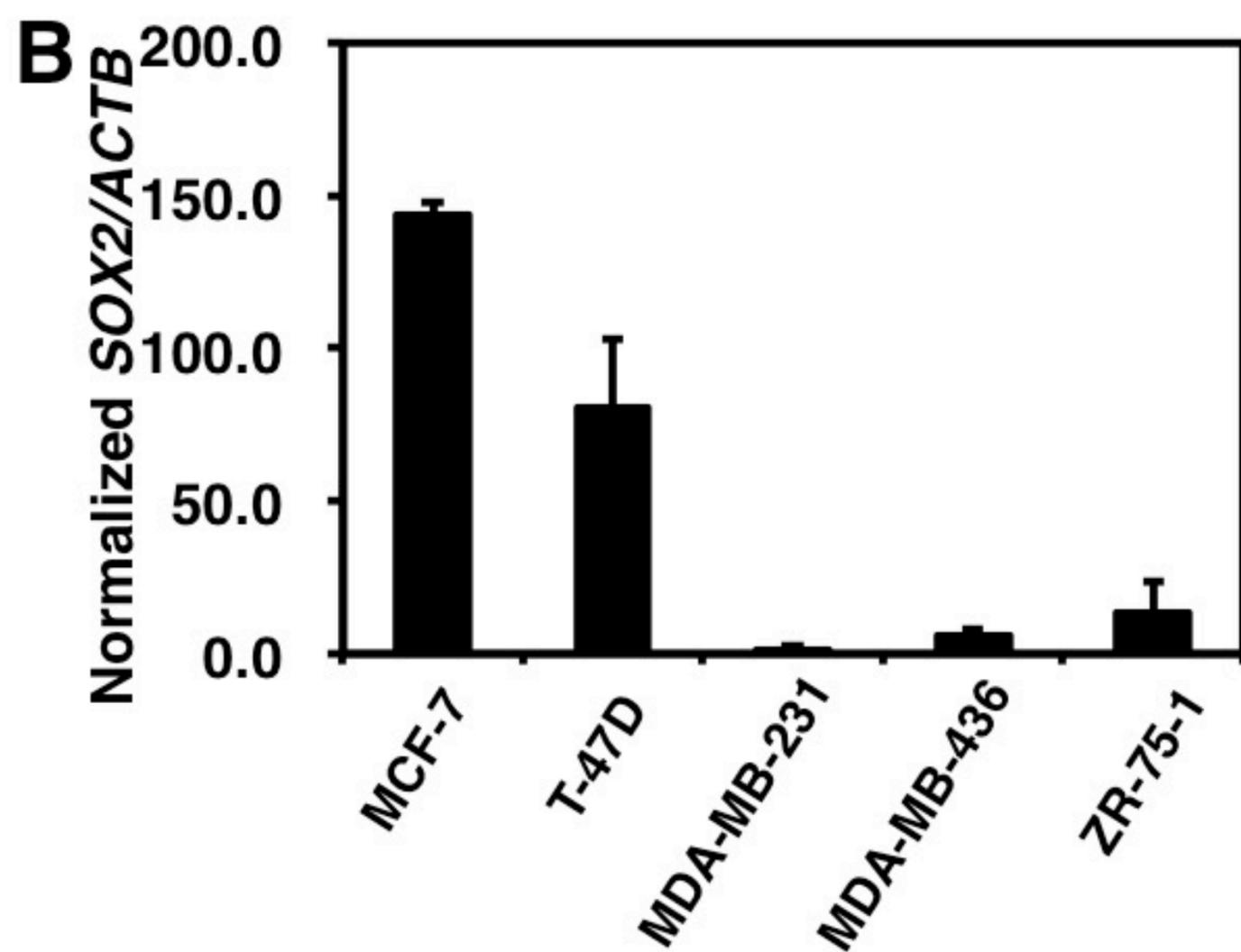
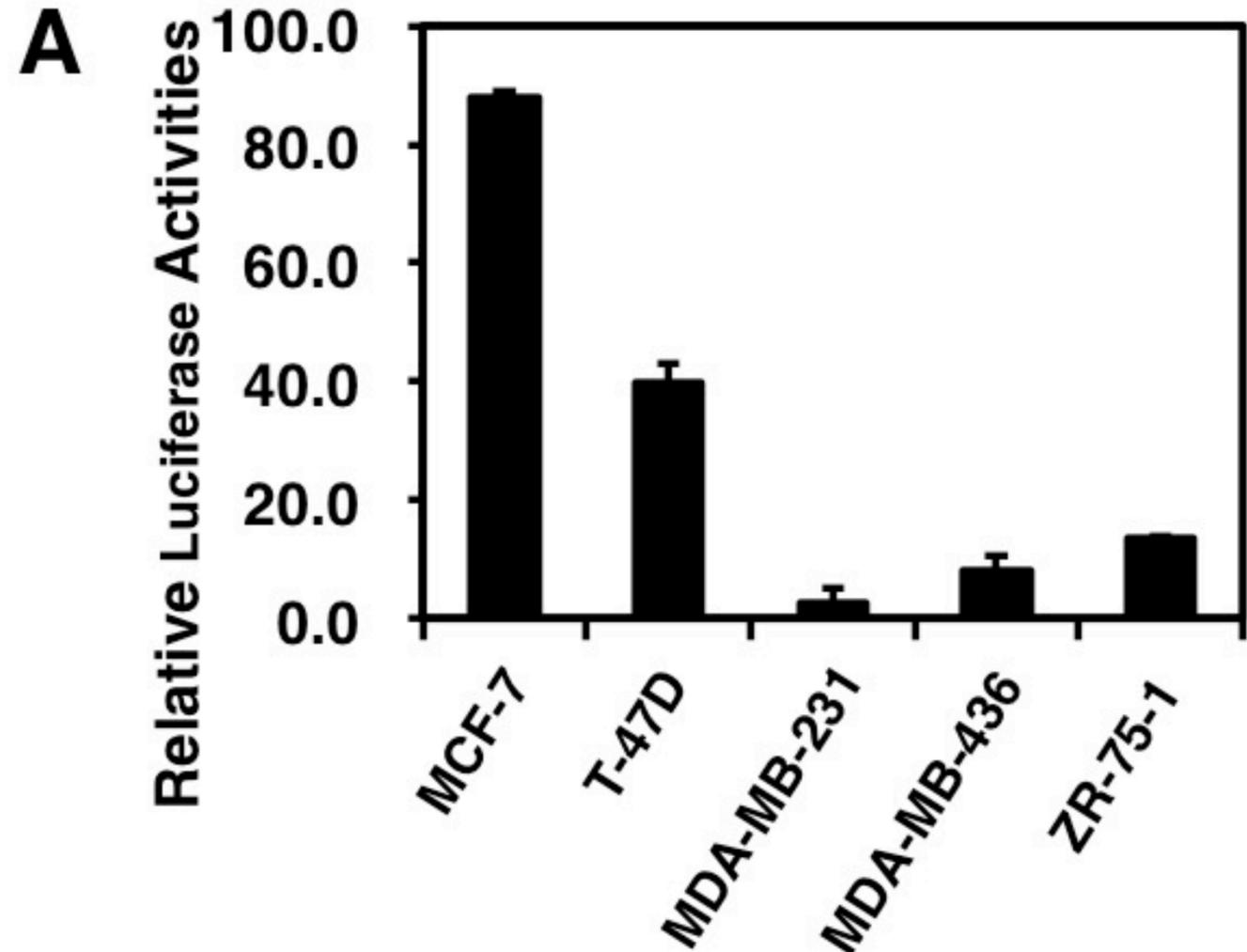
Figure legends

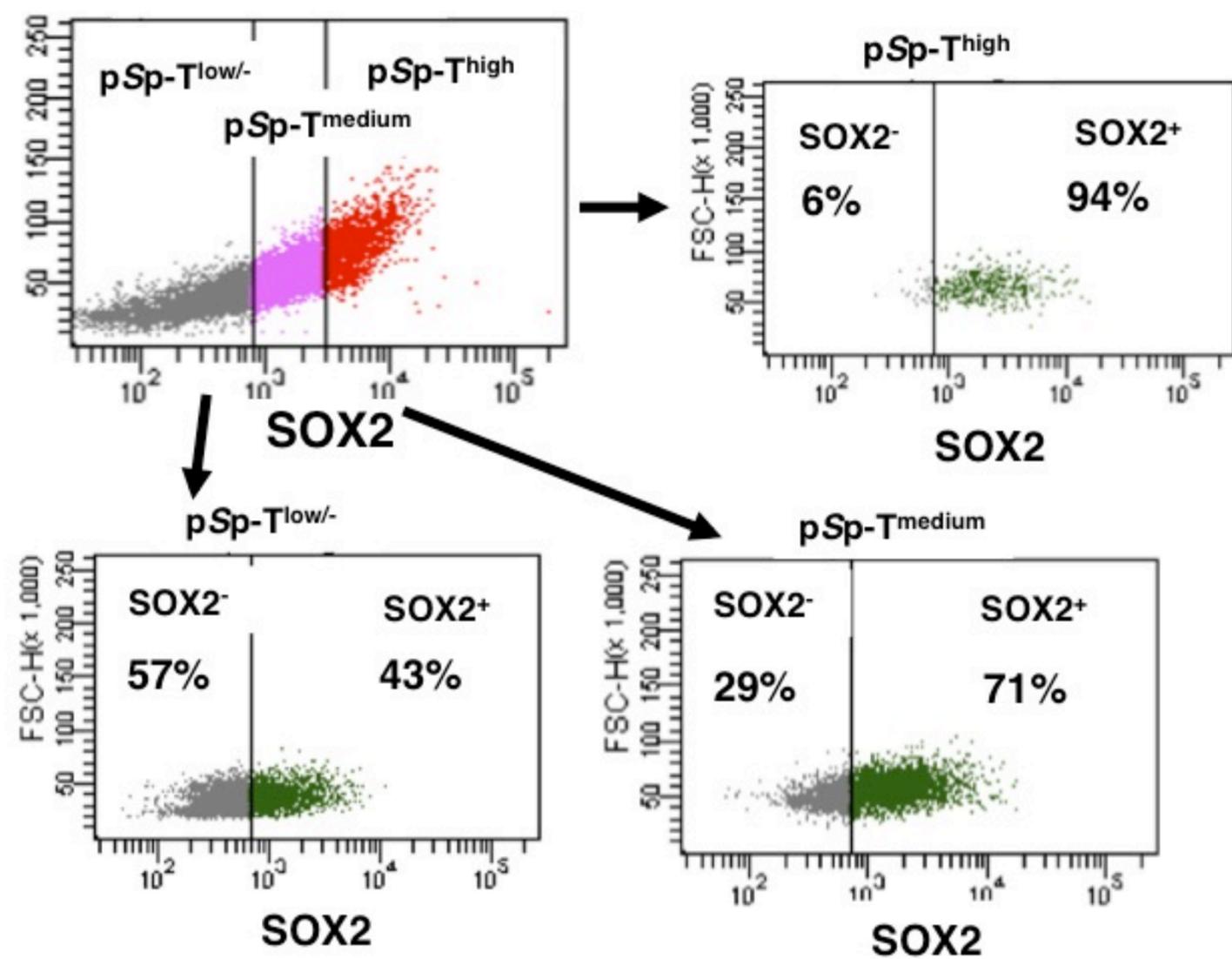
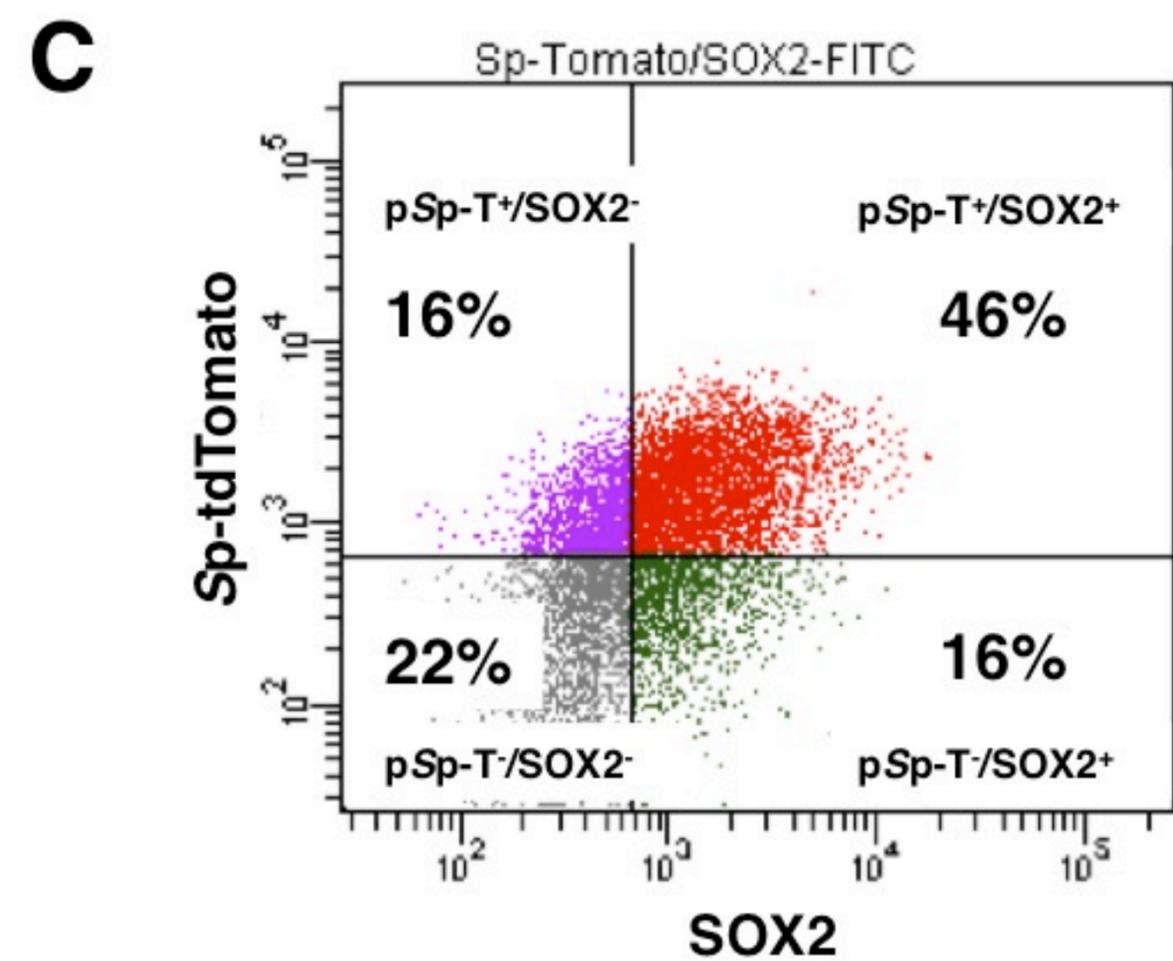
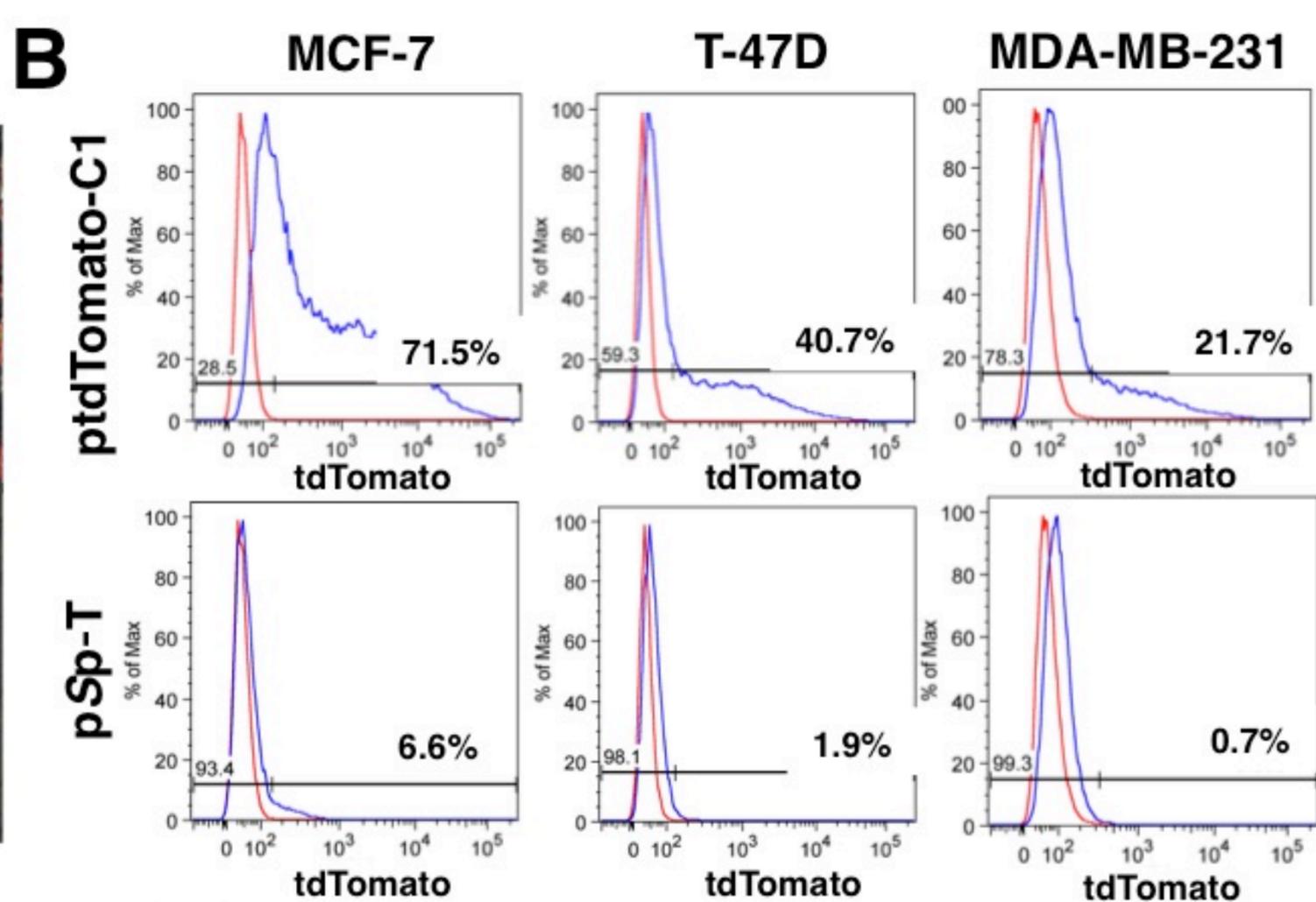
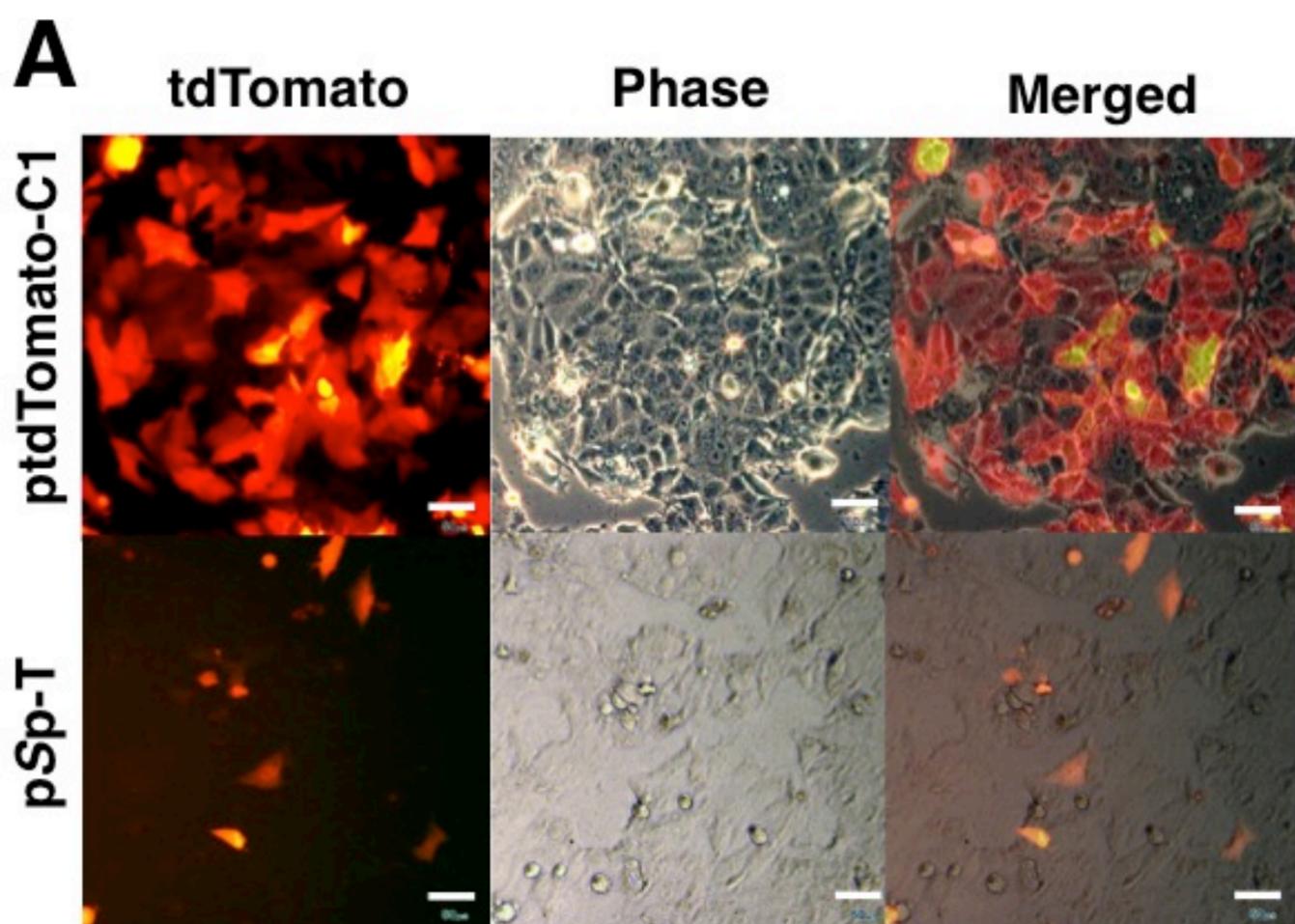
Fig. 1. *SOX2* promoter activities and expressions of endogenous *SOX2* mRNA in human breast cancer cell lines. (A) *SOX2* promoter activities in breast cancer cell lines MCF-7, T-47D, MDA-MB-231, MDA-MB-436 and ZR-75-1. The cells were co-transfected with pSp-L reporter vector and internal control vector (pRL-TK). Values from reporter firefly luciferase were normalized to those of internal control Renilla luciferase. Each column with bar represents the mean \pm SD of triplicate samples. (B) Expression levels of endogenous *SOX2* mRNA by qRT-PCR analysis. Expression levels of *SOX2* mRNA were normalized to those of β -actin (*ACTB*). Each column with bar represents the mean \pm SD of triplicate samples. (C) *SOX2* gene enhancer activities in MCF-7, T-47D and MDA-MB-231 cells. The cells were co-transfected with reporter vectors (pSp-L, pSI-Sp-L, pSp-L-SII or pSI-Sp-L-SII) and pRL-TK vector. Values from reporter firefly luciferase were normalized to those of internal control Renilla luciferase. Each column with bar represents the mean \pm SD of triplicate samples. White column: MCF-7, Black Colum: T-47D, Shaded column: MDA-MB-231. * $p < 0.05$ vs Sp. (D) Relationship between *SOX2* promoter activities and *SOX2* mRNA expression levels in the breast cancer cell lines. Pearson product moment correlation coefficient was $R^2 = 0.97894$.

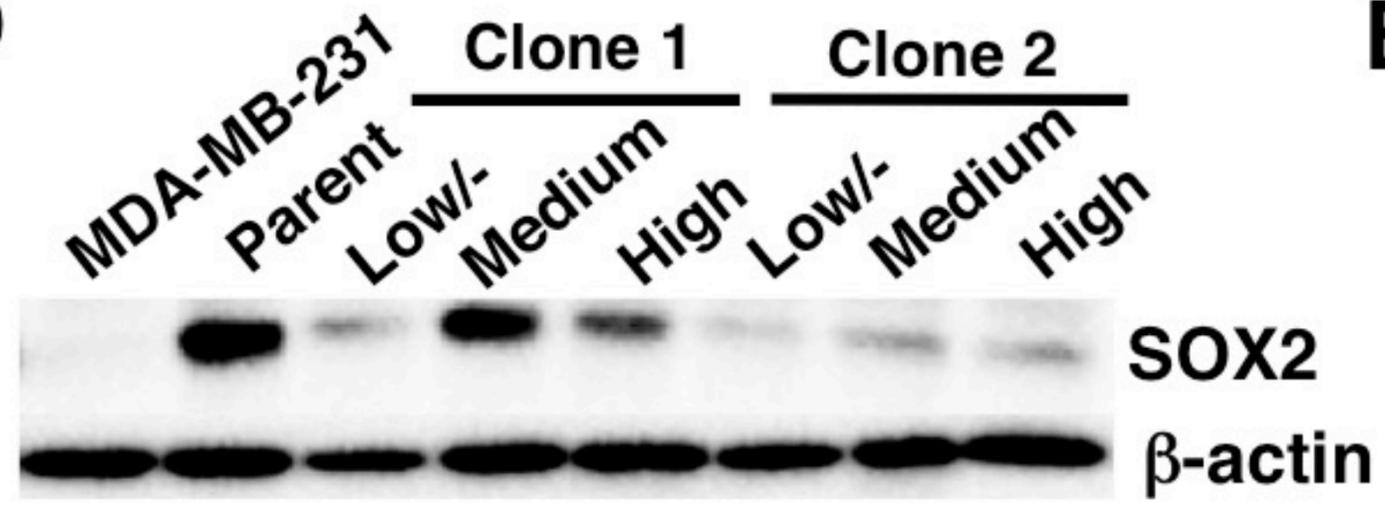
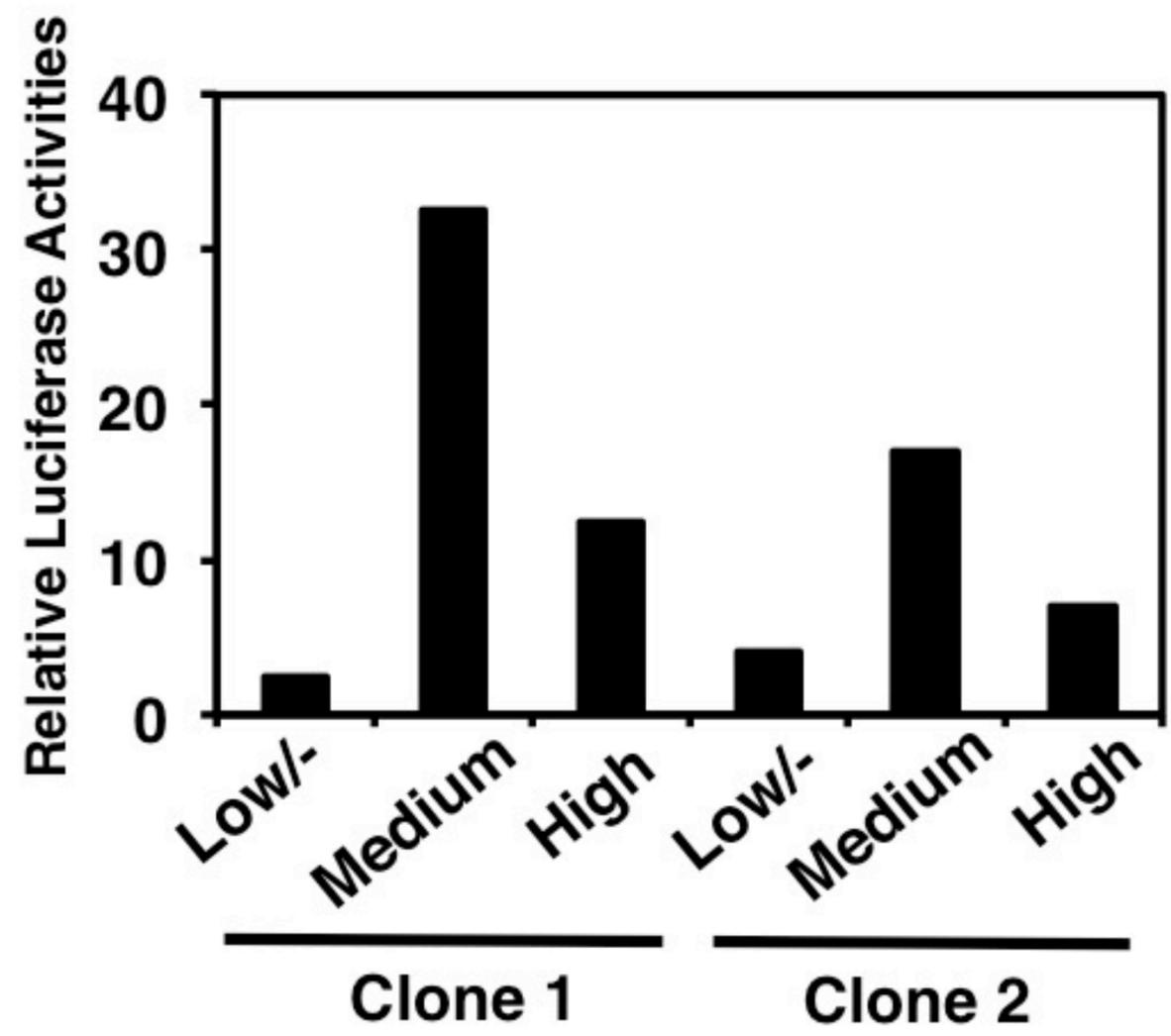
Fig. 2. Detection of the cells expressing fluorescence protein tdTomato gene driven by *SOX2* promoter. (A) Microphotographs of MCF-7 cells transfected with pSp-T or pCMV-tdTomato-C1 vector. ptdTomato-C1 was used for measuring transfection efficiency. White Bar, 50 μ m (B) FACS analysis of three human breast cancer cell lines transfected with pSp-T reporter or pCMV-tdTomato-C1 vector. Red lines: Non-transfectant, Blue lines: Transfectants with ptdTomato-C1(upper) or pSp-T (lower). (C) FACS analysis of *SOX2* protein expression in clonal MCF-7 subline transfected with pSp-T (Clone 1 cells). Clone 1 cells were fixed and immunofluorescently stained with anti-*SOX2* antibody, and were sorted into three groups on the basis of tdTomato fluorescence intensity. Left Pannel, Silver dots: pSp-T⁻/*SOX2*⁻, Purple dots: pSp-T⁺/*SOX2*⁻, Red dots: pSp-T⁺/*SOX2*⁺, Dark Green dots: pSp-T⁻/*SOX2*⁺. Right Pannel, Silver dots: pSp-T^{low/-}, Purple dots: pSp-T^{medium}, Red dots: pSp-T^{high}, Dark Green dots: *SOX2*⁺. (D) Western blotting analysis of *SOX2* protein expressions in MCF-7 derived Clone 1 and Clone 2 cells expressing tdTomato gene driven by *SOX2* promoter. (E) Activities of luciferase driven by *SOX2* promoter in the Clone 1 and Clone 2 cell populations with different expression levels of tdTomato gene driven by *SOX2* promoter.

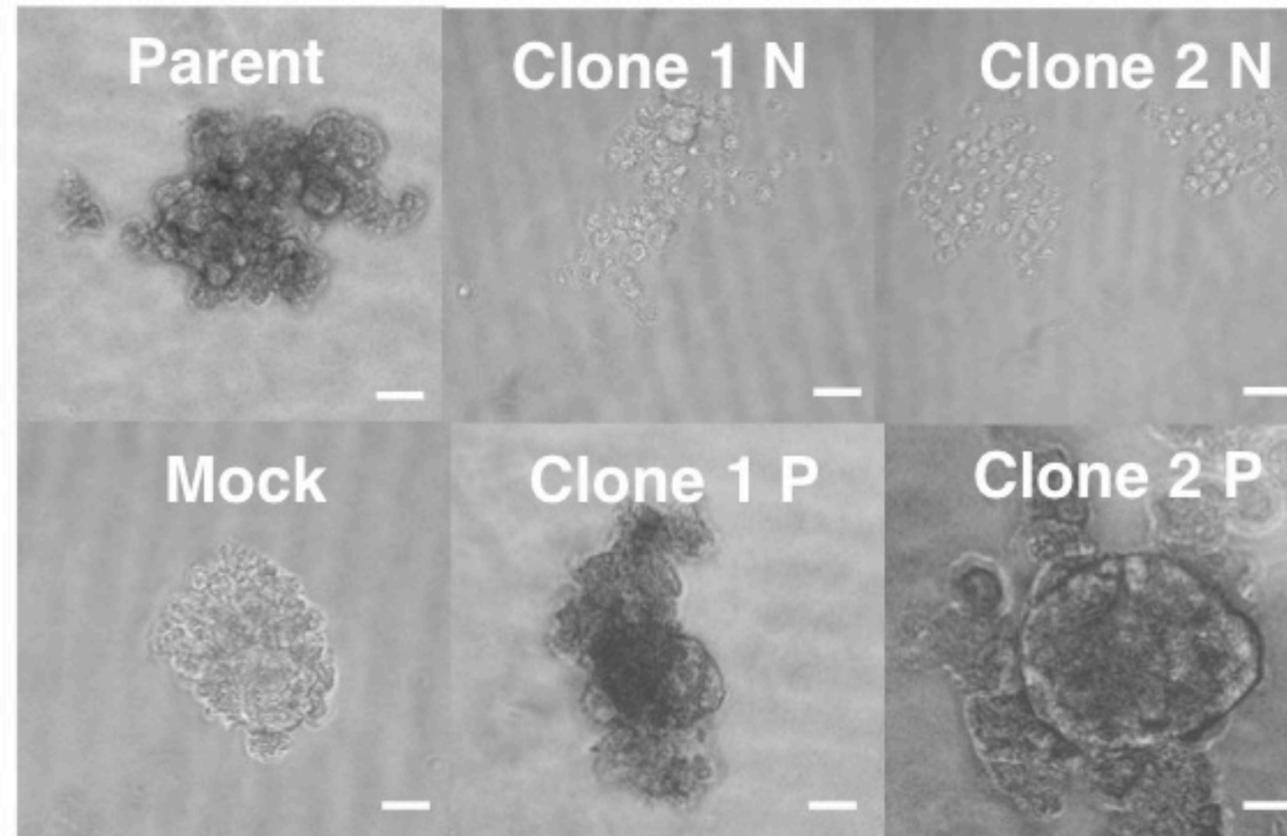
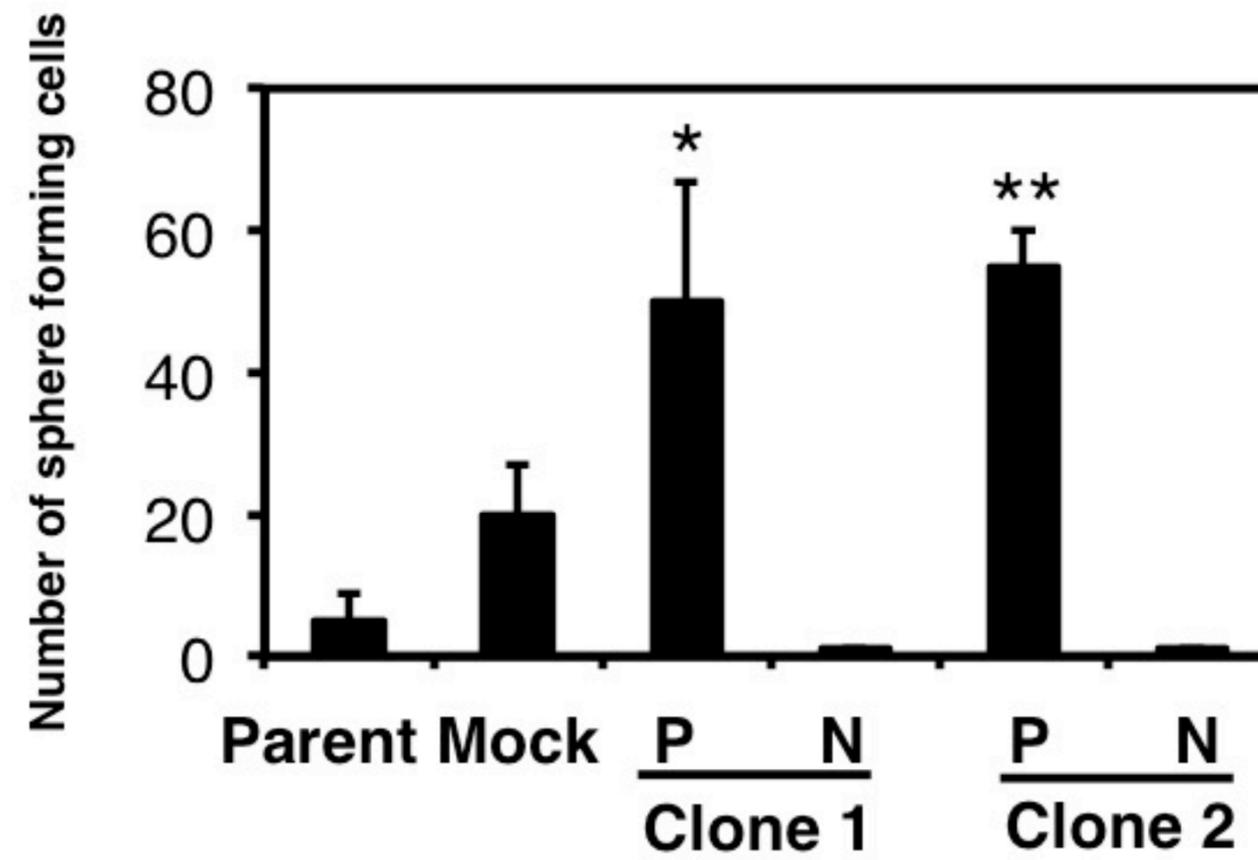
Fig. 3. Sphere-forming ability of MCF-7 Clone 1- or Clone 2-derived subpopulations with or without *SOX2* promoter activity. (A) Phase-contrast microphotographs of spheres days after seeding of MCF-7 parent cells (Parent), MCF-7 cells transfected with pCMV-tdTomato-C1 (Mock), tdTomato fluorescence-positive (P) or low/negative (N) cells derived from cloned MCF-7 sublines transfected with pSp-T vector (Clone 1 and 2). White Bar, 10 μ m. (B) Comparison of sphere-forming ability among Parent, Mock, P and N cells. Each column and bar represents the mean \pm SD of three independent experiments. * $p < 0.05$; ** $p < 0.001$ vs Parent, Mock or N.

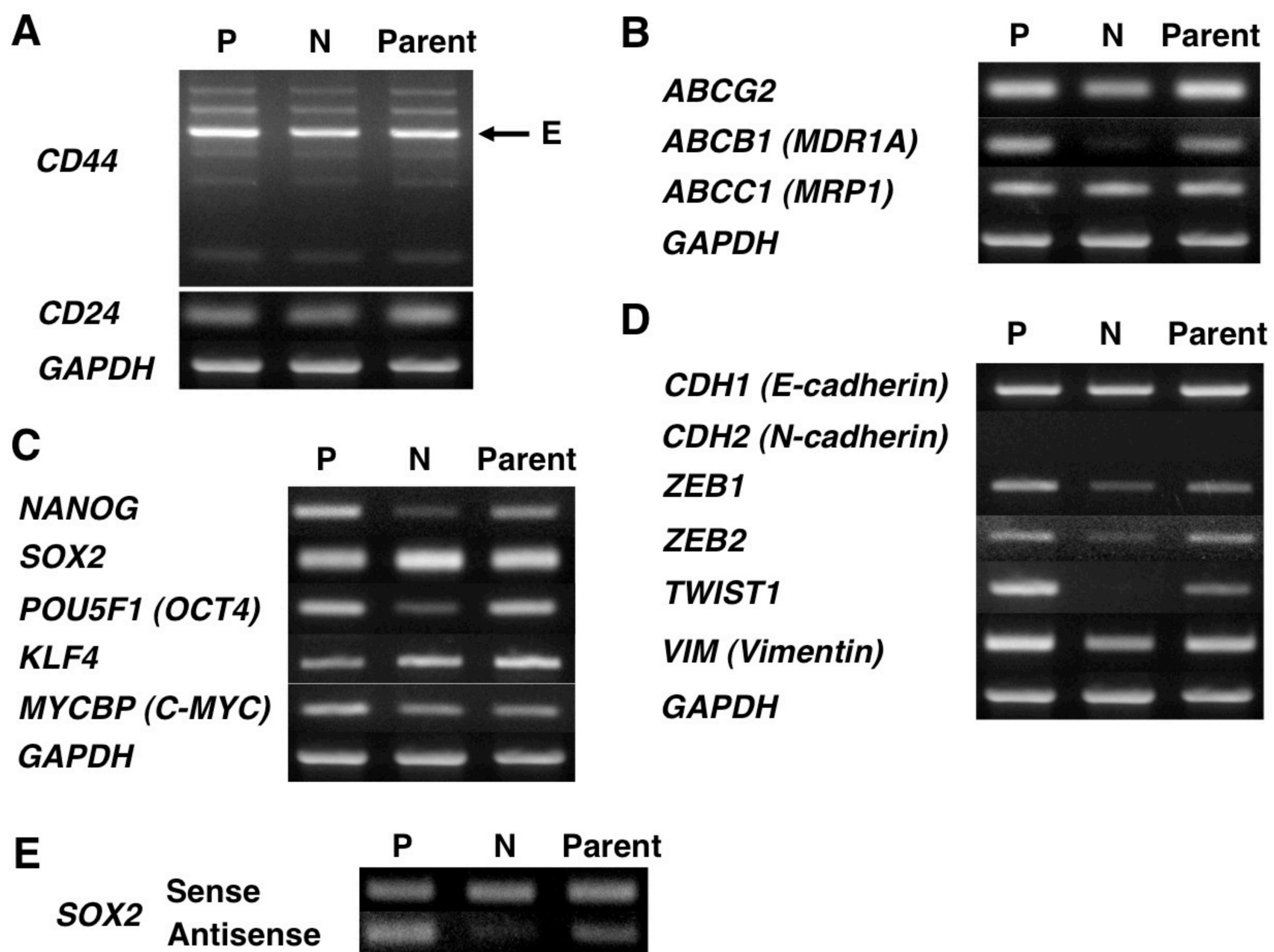
Fig. 4. mRNA expression profiles of stemness-related genes in MCF-7 cells with *SOX2* promoter activity. (A) Breast cancer stem cell marker genes. An arrow indicates bands of PCR products of epithelial-type *CD44*. (B) Side population-related genes. (C) Stemness-related transcription factor genes. (D) EMT-related marker genes. (E) *SOX2* sense and antisense. *GAPDH* was used as an internal control. P: pSp-T⁺ cells, N: pSp-T^{low/-} cells, Parent: Clone 1.





D**E**

A**B**



Supplementary figure legends

Fig. 1. The effect of *SOX2* gene enhancer in human teratoma cell line PA-1. (A) Structure of *SOX2* promoter and/or enhancer *SRRs* driven luciferase gene vectors. Sp-L: *SOX2* promoter, SI-Sp: *SRR1-SOX2* promoter, Sp-SII: *SOX2* promoter-*SRR2*, SI-Sp-SII: *SRR1-SOX2* promoter-*SRR2*. (B) *SOX2* promoter and *SRRs* activities in teratoma cell line PA-1. The cells were co-transfected with pSp-L or pSI-Sp-L or pSp-L-SII or pSI-Sp-L-SII reporter vector and internal control vector (pRL-TK). Values from reporter firefly luciferase were normalized to those of internal control Renilla luciferase. Each column with bar represents the mean \pm SD of triplicate samples. (C) Structure of HMG and POU binding sites and mutans. (D) The effect of mutation of HMG and POU binding sites in *SRR2*. The cells were co-transfected with pSp-L-SII having HM or PM and pRL-TK vector. Values from reporter firefly luciferase were normalized to those of internal control Renilla luciferase. Each column with bar represents the mean \pm SD of triplicate samples. * $p < 0.05$ vs Sp.

Fig. 2. FACS analysis of *SOX2* protein expression in clonal MCF-7 subline transfected with pSp-T (Clone 2 cells). Clone 2 cells were fixed and immunofluorescently stained with anti-*SOX2* antibody, and were sorted into three groups on the basis of tdTomato fluorescence intensity. Left Panel, Silver dots: pSp-T⁻/*SOX2*⁻, Purple dots: pSp-T⁺/*SOX2*⁻, Red dots: pSp-T⁺/*SOX2*⁺, Dark Green dots: pSp-T⁻/*SOX2*⁺. Right Panel, Silver dots: pSp-T^{low/-}, Purple dots: pSp-T^{medium}, Red dots: pSp-T^{high}, Dark Green dots: *SOX2*⁺.

