Transcriptional repressor GATA binding 1–mediated repression of SRY-box 2 expression suppresses cancer stem cell functions and tumor initiation

Xue Gong§, Weiguang Liu§, Lele Wu§, Zhifang Ma§, Yuzhi Wang§, Shiyi Yu§, Jun Zhang§, Hao Xie§, Guanyun Wei§, Fei Ma§, Ling Lu§, and Liming Chen§‡§

From the §Institute of Life Science, Southeast University, Nanjing 210096, China, the †Jiangsu Key Laboratory for Molecular and Medical Biotechnology, College of Life Science, Nanjing Normal University, Nanjing 210023, China, the ‡Department of Otolaryngology-Head and Neck Surgery, DrumTower Clinical Medical College of Nanjing Medical University, Nanjing 210008, China, and the †Department of Otolaryngology-Head and Neck Surgery, Nanjing Drum Tower Hospital Affiliated to Nanjing University Medical School, Nanjing, Jiangsu 210008, China

Edited by Xiao-Fan Wang

Cancer stem cells (CSCs) have been reported in a variety of cancers. SRY-box 2 (SOX2) is a member of the SOX family of transcription factors and has been shown to play a critical role in maintaining the functions of CSCs and promoting tumor initiation. However, the underlying mechanisms for the transcriptional regulation of the SOX2 gene in CSCs are unclear. In this study, using in silico and experimental approaches, we identified transcriptional repressor GATA binding 1 (TRPS1), an atypical GATA-type transcription factor, as a critical transcriptional regulator that represses SOX2 expression and thereby suppresses cancer stemness and tumorigenesis. Mechanistically, TRPS1 repressed SOX2 expression by directly targeting the consensus GATA-binding element in the SOX2 promoter as elucidated by ChIP and luciferase reporter assays. Of note, in vitro mammosphere formation assays in culture and in vivo xenograft tumor initiation experiments in mouse models revealed that TRPS1-mediated repression of SOX2 expression suppresses CSC functions and tumor initiation. Taken together, our study provides detailed mechanistic insights into CSC functions and tumor initiation by the TRPS1–SOX2 axis.

Results and discussion

TRPS1 represses SOX2 expression

SOX2 plays an important role in CSC functions (18). To address how SOX2 expression is regulated, we set out to identify new candidate transcription factors targeting SOX2. By analyzing the promoter sequence of SOX2 in silico, we determined that SOX2 promoter contains the WGATAR sequence element, which is a well characterized consensus sequence recognized by GATA transcription factors (Fig. 1A). The family of GATA factors is composed of GATA1–6 and an atypical GATA-type zinc finger motif. TRPS1 is implicated in stem cell functions because its overexpression promotes differentiation of chondrogenic cells (12), and TRPS1 regulates the Wnt signaling pathway in vibrissa follicle morphogenesis (13). TRPS1 expression is also elevated in luminal breast cancer as compared with basal breast cancer (14), which contains more CSC population than luminal breast cancer (15). Mutations in TRPS1 have been reported in basal-like breast cancer (16, 17). However, whether and how TRPS1 contributes to CSC functions and tumor initiation by regulating SOX2 expression is mostly unknown. In the current study, we report that TRPS1 represses SOX2 expression by directly targeting its promoter and thus suppressing CSC functions and tumor initiation.

This work was supported by National Natural Science Foundation of China Grants 81772956 and 81572712, National Basic Research Program of China 973 Program Grant 2015CB965000, Key University Science Research Project of Jiangsu Province Grant 17KJA320002, Natural Science Foundation of Jiangsu Province Grant BK2016030027, funds from the Priority Academic Program and the Development of Jiangsu Higher Education Institutions, Jiangsu Provincial Medical Youth Talent of the Project of Invigorating Health Care through Science, Technology and Education Grant QNRC2016002, and National Science Foundation for Young Scientists of China Grant 81700913. The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Table S1 and Figs. S1–S7.

References

1. To whom correspondence should be addressed. Tel.: 86-25-85891050; Fax: 86-25-85891050; E-mail: chenliming1981@njnu.edu.cn
2. The abbreviations used are: CSC, cancer stem cell; TRPS1, transcriptional repressor GATA binding 1; SOX2, SRY-box 2; GEO, Gene Expression Omnibus; ANOVA, analysis of variance; PE, R-phycoerythrin; APC, Allophycocyanin.

Cancer stem cells (CSCs) are at the root of tumor recurrence and metastasis in many human cancers (1). Numerous studies have shown the important role of SRY-box 2 (SOX2) in CSC functions in many cancer types, including invasive skin squamous-cell carcinoma (2), melanoma (3), glioma (4), and breast cancer (5, 6). However, regulation of SOX2 expression remains unclear.

GATA factors play important roles both in development and cancer. Typical GATA factors (GATA1–6) are implicated in inducing differentiation of embryonic stem cells (7–9) and modulating CSC functions (10, 11). Transcriptional repressor GATA binding 1 (TRPS1), an atypical member of GATA transcriptional factor family, contains an atypical GATA-type zinc finger motif. TRPS1 is implicated in stem cell functions because its overexpression promotes differentiation of chondrogenic cells (12), and TRPS1 regulates the Wnt signaling pathway in vibrissa follicle morphogenesis (13). TRPS1 expression is also elevated in luminal breast cancer as compared with basal breast cancer (14), which contains more CSC population than luminal breast cancer (15). Mutations in TRPS1 have been reported in basal-like breast cancer (16, 17). However, whether and how TRPS1 contributes to CSC functions and tumor initiation by regulating SOX2 expression is mostly unknown. In the current study, we report that TRPS1 represses SOX2 expression by directly targeting its promoter and thus suppressing CSC functions and tumor initiation.
tion with the SOX2 expression, suggesting that TRPS1 represses SOX2 expression (Fig. 1, B and C, and Fig. S1). This observation was consistent with the notion that TRPS1 has intrinsic transcriptional repression activity (22). Also, elevated expression of TRPS1 was observed in highly differentiated luminal breast cancers compared with less differentiated basal breast cancers (14). SOX2 is believed to be a possible driver of the basal-like phenotype in sporadic breast cancers (23). Based on these observations, we hypothesized that TRPS1 is a repressive factor for SOX2 expression and suppresses CSC functions. To test this hypothesis, which fits well with the current knowledge on TRPS1 and SOX2, we attempted to experimentally confirm that TRPS1 negatively regulates SOX2 expression. We silenced TRPS1 in T47D and MCF7 cells, which are breast cancer cell lines with elevated endogenous TRPS1 expression, and found that silencing TRPS1 led to up-regulation of SOX2 at both mRNA and protein levels (Fig. 1D). These results indicate that TRPS1 represses SOX2 expression.

**TRPS1 targets SOX2 promoter**

To verify whether TRPS1 represses SOX2 expression by directly targeting the GATA-binding element in the SOX2 promoter, we performed the ChIP–quantitative PCR assay using TRPS1 antibodies in T47D and MCF7 cells. The results showed significant enrichment of TRPS1 at the SOX2 promoter (>5-fold enrichment) in both T47D and MCF7 cells using a well characterized untranscribed genomic region, Untr4, as a control (Fig. 2A). Furthermore, we carried out luciferase reporter

---

**Figure 1. TRPS1 represses SOX2 expression.** A, schematic diagram of the SOX2 gene with the GATA elements highlighted in its promoter region. B, expression of GATA factors in breast cancers from GEO. C, correlation between GATA factors and SOX2 in breast cancers from GEO. r, correlation coefficient. D, silencing TRPS1 increases SOX2 expression at both mRNA and protein level. All data are presented as the means ± S.D. of three independent experiments. The t test was used for statistic quantification. ***, p < 0.001.
assay and observed that ectopic overexpression of TRPS1 could significantly reduce the luciferase activity of the WT SOX2 promoter (WT–SOX2 promoter) and Mut1–SOX2 promoter with a mutation in the noncore TRPS1-binding element in the SOX2 promoter. The reduction in luciferase activity with overexpressed TRPS1 was lost in Mut2–SOX2 promoter with a mutation in the core TRPS1-binding consensus element GATA in the SOX2 promoter (Fig. 2B). Also, silencing TRPS1 significantly increased the luciferase activity of the SOX2 promoter in MCF7 cells (Fig. 2C). These observations indicated that TRPS1 repressed transcriptional activity of SOX2 promoter by targeting its GATA-binding element. Domain structure analysis revealed that TRPS1 contained three domains: a GATA-binding domain flanked with N-terminal and C-terminal domains.

To further explore which domain in TRPS1 was important for its repressive function on SOX2 expression, we constructed several TRPS1 truncates and carried out luciferase reporter assay. The results showed that only full-length TRPS1 and a
TRPS1 truncate with C-terminal domain together with the GATA domain were able to repress the transcriptional activity of the SOX2 promoter, indicating that C-terminal and GATA domains of TRPS1 were required for the TRPS1 transcriptional repression function on the SOX2 promoter (Fig. 2D). Taken together, these results indicated that TRPS1 repressed SOX2 expression by directly targeting the GATA element in the SOX2 promoter.

**TRPS1 suppresses CSC functions by repressing SOX2 expression in vitro**

SOX2 is a key stem cell marker and has been shown to promote tumor initiation (2, 5, 24). As described above, TRPS1 targets SOX2 promoter to repress its expression (Figs. 1 and 2). We therefore proposed that TRPS1 regulates CSC functions by repressing SOX2 expression. In addition to SOX2, other documented stem cell pluripotency markers include OCT4, KLF4, and NANOG (25–27). We investigated whether OCT4, KLF4, and NANOG were involved in the regulation of stemness of cancer cells by TRPS1. The expression of OCT4, KLF4, and NANOG upon TRPS1 silencing was analyzed and compared with that of SOX2. The results showed that only expression of SOX2 and not OCT4, KLF4, and NANOG was significantly up-regulated upon silencing TRPS1 (Fig. 3A and Fig. S2).

To test whether the TRPS1–SOX2 axis contributed to CSC functions, we carried out an *in vitro* mammosphere formation assay considered to be the gold standard for testing CSC functions *in vitro* (28, 29). TRPS1 was successfully knocked down using specific anti-TRPS1 shRNA, and the increased SOX2 upon TRPS1 knockdown could be reversed by transfecting with additional anti-SOX2 shRNA (Fig. 3B). Moreover, overexpression of a shRNA-resistant TRPS1 rescued the SOX2 up-regulation caused by the TRPS1 knockdown (Fig. 3B). In a mammosphere assay, the results showed that silencing TRPS1 in MCF7 cells increased the number of mammospheres, whereas silencing SOX2 had an opposite effect. As expected, additional silencing of SOX2 in cells with silenced TRPS1 fully restored the increased mammosphere formation ability of cells with silencing TRPS1 alone (Fig. 3C). These observations indicated that TRPS1 suppressed the sphere-forming ability of cancer cells by reducing the expression of SOX2. To further confirm the effects of the TRPS1–SOX2 axis on CSC functions, the CD44+/CD24− cell population, which is reported to have CSC properties, was investigated using flow cytometry. The results showed that silencing TRPS1 in MCF7 cells led to a significant increase in the CSC population (CD44+/CD24−) (up to ~3-fold), whereas silencing SOX2 decreased the CSC population, and additional silencing SOX2 in cells with silenced TRPS1 fully restored the increased CSC population (Fig. 3D and Fig. S3). Also, CSCs isolated from MCF7 cells using FACS showed lower expression of TRPS1 and higher expression of SOX2 compared with those in other non-CSCs (Fig. S4). Consistently, highly purified CSCs (ALDH+/CD44+/CD24−) from breast cancer patient-derived xenografts expressed higher SOX2 and lower TRPS1 compared with other types of breast cancer cells (Fig. S5). These results provided further evidence that TRPS1 might suppress CSC function by repressing SOX2 expression *in vitro*.

**TRPS1 suppresses tumor initiation by repressing SOX2 expression in vivo**

Tumor initiation ability with the transplantation of limiting dilutions of cancer cells is a direct and convincing evidence for CSC functions (30, 31). We used xenograft tumor initiation models to validate that TRPS1 suppressed CSC functions by repressing SOX2 expression. Female SCID mice were subcutaneously injected with limiting dilutions of MCF7 cells (5 × 10⁶, 5 × 10⁵, and 5 × 10⁴ cells) with silencing TRPS1 alone or with additional silencing of SOX2. The results showed that compared with mice treated with control shRNA, silencing TRPS1 resulted in decreased tumor latency and this phenotype was rescued by additionally silencing SOX2 (Fig. 4A). Furthermore, cells with silenced TRPS1 had a much greater ability and cells with additional silencing of SOX2 had a much smaller ability to form tumors as compared with controls (Fig. 4, A and B, and Table 1). These results suggested that TRPS1 inhibited CSC functions and suppressed breast tumor initiation by repressing SOX2.

We examined tumor growth in xenograft tumor models at two cell densities, 5 × 10⁶ and 5 × 10⁵, over a period of 6 weeks. We observed that silencing of TRPS1 suppressed tumor growth, which may indicate a different role of TRPS1 in the tumor promotion stage of oncogenesis (Fig. 4, C and D). Also, additional silencing of SOX2 with silencing TRPS1 further suppressed tumor growth compared with silencing with TRPS1 alone (Fig. 4, C and D). This observation was likely due to the delayed tumor formation upon silencing SOX2, which led to relatively short tumor growth time in xenograft mouse models. These discrepant results raise an interesting possibility of differential roles of TRPS1 in tumor initiation versus tumor promotion. These apparently conflicting results can be explained, at least in part, by the current notion of the proliferation ability of CSCs, which are critical for tumor initiation but show slow proliferation rate and self-renewal capabilities (32, 33). The TRPS1 expression is believed to have a dynamic role in the process of tumor development (34). In this respect, we recently reported that TRPS1 promoted tumor growth by functioning as a scaffold to regulate USP4-directed HDAC2 deubiquitination. These observations provided a mechanistic insight into the tumor growth-promoting role of TRPS1 (35).

In addition to TRPS1, GATA3 was the other GATA factor that was found to be frequently expressed in breast cancer. Consistent with the observation that GATA3 promoted luminal breast cancer cell differentiation (36, 37), we found that silencing GATA3 led to increased expression of SOX2 and decreased expression of TRPS1 (Fig. S6). Also, silencing both GATA3 and TRPS1 in MCF7 cells increased expression of SOX2 (Fig. S6). At this point, it is difficult to conclude whether GATA3 and TRPS1 compete or cooperate with each other in modulating SOX2 expression and CSC functions. However, our results, together with the fact that both GATA3 and TRPS1 recognize the GATA element, indicated that depending on the cellular context, both GATA3 and TRPS1 might target SOX2 to regulate SOX2 expression and CSC functions. It would be worthwhile to further explore the precise mechanism in the future.
In conclusion, our current study has shown that TRPS1 recognizes SOX2 promoter and represses SOX2 expression, thereby suppressing CSC functions. Experimental procedures

Cell culture

Cell lines T47D, MCF7, and 293T were kind gifts from Drs. Zhao Bin, Zhu Tao, and Lin Chenqi. Cell lines T47D and MCF7 were maintained in RPMI 1640 (Wisent, catalog no. 350-000-CL) with 10% fetal bovine serum (HyClone, catalog no. 10010-044).
SV30087.02) and 1% penicillin/streptomycin (Wisent, catalog no. 450-201-EL). The 293T cells were cultured under standard conditions in Dulbecco’s modified Eagle’s medium (Wisent, catalog no. 319-005-CL), 10% fetal bovine serum, 1% penicillin/streptomycin. All three cell lines were cultured at 37 °C in a 5% CO2 incubator and used within 6 months. The cell lines were authenticated by the short tandem repeat typing. Trypsin/EDTA and PBS were purchased from Wisent Inc.

RNAi

The siRNA transfections were carried out with Lipofectamine RNAi-MAX (Thermo Fisher, catalog no. 13778100) in the antibiotic-free medium according to the manufacturer’s instructions. The sequences for control siRNA was the sense (5'-UUC UCC GAA CGU GUC UGU UTT-3') and antisense (5'-ACG UGA CAC GGU UCG AGA ATT-3'). The sequences for SOX2 siRNA were: 1) sense (5'-GGUUGACAC CGU UGG UAA UTT-3') and antisense (5'-ACG UGA CAC GUU CGG AGA ATT-3'); and 2) sense (5'-UGC CGA GAA UCC AUG UAU ATT-3') and antisense (5'-UGU GCC UUG ACC GUC GGC ATT-3'). The sequences for TRPS1 siRNA were: 1) sense (5'-GUC CCU UGA AUG UAG UAA ATT-3') and antisense (5'-UUU ACU ACA UUC AAG GGA CTT-3'); and 2) sense (5'-GCA CAC AGC UGC UAC AAA UTT-3') and antisense (5'-AUU UGU AGC AGC UGU GUG CTT-3'). The sequences for GATA3 siRNA were: 1) sense (5'-CAT CGA CGG TCA AGG CAA CTT-3') and antisense (5'-GUU GCC UUG ACC GUG GAA GGT-3'); and 2) sense (5'-AAC AUC GAC GGU CAA GGC AAA-3') and antisense (5'-GUU GCC UUG ACC GUG GUA-3').

Gene-expression analysis by quantitative real-time PCR and Western blotting

Total RNA was prepared with E.Z.N.A.® total RNA kit I (OMEGA Bio-tek, catalog no. R6834-02), and 2 μg of total RNA was transcribed into cDNA with PrimeScript™ RT reagent kit (Takara Bio, catalog no. RR036A). RT-PCR was carried out as previously described (38) using specific primers (Table S1). β-Actin was used as an internal control.

Cell lysates were prepared as previously described (38). The antibodies used in the Western blots were SOX2 (1:1000, R&D Systems, catalog no. AF2018), TRPS1 (1:1000, R&D Systems, catalog no. AF4838), GATA3 (1:1000, Abcam, catalog no. ab199428), rabbit anti-DDDDK-Tag pAb (1:2500, Abclonal,

Table 1

| Tumor incidence of MCF7 cells expressing indicated shRNAs |
|----------------------------------------------------------|
| Number of cells  | shControl | shTRPS1 | shTRPS1 + shSOX2 |
| 5 x 10^6         | 7/8       | 8/8     | 3/8             |
| 5 x 10^5         | 4/8       | 6/8     | 1/8             |
| 5 x 10^4         | 1/8       | 5/8     | 0/8             |

*p value vs. shControl
*p value vs. shTRPS1

The *p value was obtained by Pearson’s chi-squared test using ELDA software (40).

Figure 4. TRPS1 suppresses tumor initiation ability by repressing SOX2 expression. A, Kaplan–Meier curves showing the palpable tumor-free survival of mouse xenografts after subcutaneous injection with the indicated number of MCF-7 cells. The *p values were calculated with a log-rank test. B, percentage of palpable tumor-free mice with subcutaneous injection with the limited number of diluted cells upon silencing TRPS1, SOX2, or both. C, palpable tumor growth across time. D, tumor volumes at the end point (week 6). Error bars represent means ± S.D. One-way ANOVA followed by Newman–Keuls test was used for statistic quantifications. **, *p < 0.01; ***, *p < 0.001; ns, not significant.

SV30087.02 and 1% penicillin/streptomycin (Wisent, catalog no. 450-201-EL). The 293T cells were cultured under standard conditions in Dulbecco’s modified Eagle’s medium (Wisent, catalog no. 319-005-CL), 10% fetal bovine serum, 1% penicillin/streptomycin. All three cell lines were cultured at 37 °C in a 5% CO2 incubator and used within 6 months. The cell lines were authenticated by the short tandem repeat typing. Trypsin/EDTA and PBS were purchased from Wisent Inc.

RNAi

The siRNA transfections were carried out with Lipofectamine RNAi-MAX (Thermo Fisher, catalog no. 13778100) in the antibiotic-free medium according to the manufacturer’s instructions. The sequences for control siRNA was the sense (5'-UUC UCC GAA CGU GUC UGU UTT-3') and antisense (5'-ACG UGA CAC GGU UCG AGA ATT-3'). The sequences for SOX2 siRNA were: 1) sense (5'-GGUUGACAC CGU UGG UAA UTT-3') and antisense (5'-AUU ACC ACA ACG GCC ATT-3'); and 2) sense (5'-UGC CGA GAA UCC AUG UAU ATT-3') and antisense (5'-UUU ACA ACA UUC AAG GGA CTT-3'); and 2) sense (5'-GCA CAC AGC UGC UAC AAA UTT-3') and antisense (5'-AUU UGU AGC AGC UGU GUG CTT-3'). The sequences for TRPS1 siRNA were: 1) sense (5'-GUU GCC UUG ACC GUC GGA GGT-3'); and 2) sense (5'-AAC AUC GAC GGU CAA GGC AAC-3') and antisense (5'-GUU GCC UUG ACC GUC GAU-3').

Gene-expression analysis by quantitative real-time PCR and Western blotting

Total RNA was prepared with E.Z.N.A.® total RNA kit I (OMEGA Bio-tek, catalog no. R6834-02), and 2 μg of total RNA was transcribed into cDNA with PrimeScript™ RT reagent kit (Takara Bio, catalog no. RR036A). RT-PCR was carried out as previously described (38) using specific primers (Table S1). β-Actin was used as an internal control.

Cell lysates were prepared as previously described (38). The antibodies used in the Western blots were SOX2 (1:1000, R&D Systems, catalog no. AF2018), TRPS1 (1:1000, R&D Systems, catalog no. AF4838), GATA3 (1:1000, Abcam, catalog no. ab199428), rabbit anti-DDDDK-Tag pAb (1:2500, Abclonal,
**TRPS1 suppresses cancer stem cell functions**

catalog no. AE004), and β-actin (1:5000, ProteinTech, catalog no. 60008–1-lg).

**ChiP**

ChiP assay was performed according to the published method (39). T47D cells were treated with the siRNAs for 3 days to knock down TRPS1 expression. For each ChiP assay, 30 μl of Dynabeads® (Thermo Fisher) and 3 μg of TRPS1 antibody were incubated with 50 μl of cross-linked and sonicated chromatin. ChiP DNA was purified with a PCR cleanup kit (Axygen) and measured for enrichment by quantitative real-time PCR. The primer sequences for SOX2 were sense (5’-GGG GAG TGA TTA TGG GAA GA-3’) and antisense (5’-CCC TGG TCT ACC CTT ACT CA-3’), and the primer sequences for Untr4 were sense (5’-CTC CCT GTG CTT CTC AG-3’) and antisense (5’-AAT GAA CGT GTC TCC CAG AA-3’).

**Luciferase reporter assay**

The WT and mutant SOX2 reporter constructs were generated by ligating synthetic sequences into pGL3-Basic. 293T cells were seeded in 24-well plates in triplicates, and plasmid transfection was performed when cells reached 80% confluency. Luciferase activity was measured after 48 h using the Dual-Luciferase reporter system (Promega, catalog no. E1910) according to the manufacturer’s instructions.

**Lentivirus plasmid construction and cell transfection**

To obtain stable knockdown clones for TRPS1, SOX2, and the control scrambled shRNA, sequences were cloned into the pLKO.1 plasmid with AgeI and EcoRI restriction enzymes (New England Biolabs, AgeI catalog no. R0552L, EcoRI catalog no. R0101L). The sense strand of the most efficient TRPS1-shRNA sequence was 5’-CCG GCC ACA CAG CTG CTA CAA ATG CCT CGA GGC ATT TGT AGC AGC TGT GTG CGT TTT G-3’. The sense strand of the shRNA sequence against SOX2 was 5’-CCG GCC CTG CTC ATG AAG AAG GAT AAG TCT CG AGA CTT ATC CTT CCG CAT GAG CGT TTT TG-3’. The sense strand of the control shRNA sequence was 5’-CCG GTC AGG CTA TGA AGA GAT ACC TCG AGG TAT CTC TTC ATA GCC TTA TTG TTG-3’.

293T cells were transfected with the lentivirus vector and packaging plasmid mixes. Following 48 h of transfection, the viruses were produced and used for transfection into MCF7 cells. Following 12–16 h of incubation, the viruses were removed and replaced with fresh RPMI 1640. Stable MCF7 cells were selected with 2 μg/ml puromycin (InvivoGen, catalog no. ant-pr-1) for 10 days. To rescue the knockdown phenotype, shRNA-resistant lines of TRPS1 mutant were generated from the MCF7 shRNA stable cell lines. For the TRPS1 mutant isoform, two silent mutations were introduced in the following sequence regions: A2521G and T2527A.

**Mammosphere assay**

MCF7 cells transfected with control shRNA, TRPS1 shRNA, SOX2 shRNA, or both TRPS1 shRNA and SOX2 shRNA were cultured at a density of 1,000 cells/ml in suspension in serum-free Dulbecco’s modified Eagle’s medium/F12 medium (Gibco, catalog no. 12400–024) supplemented with 20 ng/ml basic fibroblast growth factor (R&D Systems, catalog no. 133-FB-025), 20 ng/ml epidermal growth factor (PeproTech, catalog no. AF-100-15), insulin–transferrin–selenium (Sigma–Aldrich, catalog no. I3146), and B-27 supplement (Gibco, catalog no. 17504–044). The medium contained 0.5% methylcellulose (R&D Systems, catalog no. HSC001) for cell aggregation. Primary mammospheres were harvested and dissociated to single cells, counted, and reseeded for the second round of mammosphere formation. The colony numbers were counted manually under a microscope after 10 days of culturing. Numbers of mammospheres compared with the controls were represented as the means ± S.D. of three independent runs. The scale bar indicates 100 μm.

**Flow cytometric analysis**

MCF7 cells were resuspended at 1 × 10⁶ cells/100 μl of staining buffer (1 × PBS with 0.5% BSA) and incubated with primary antibodies for 30 min at 4 °C in the dark. The primary antibodies were PE mouse anti-human CD24 (catalog no. 555428), APC mouse anti-human CD44 (catalog no. 559942), PE mouse IgG2a κ isotype control (catalog no. 559319), and APC mouse IgG2b κ isotype control (catalog no. 555745) (20 μl/test, BD Pharmingen). Three control groups were established for the first sorting: 1) cells labeled with the isotype antibodies of the above two antibodies, 2) cells labeled with the anti-CD24–APC antibody and the isotype control antibody for CD24, and 3) cells labeled with the anti-CD44–PE antibody and the isotype control antibody for CD44. The cells were washed with staining buffer and centrifuged at 800 × g for 5 min. For flow cytometric analysis, the cells were resuspended in staining buffer. MCF7 cells were sorted using the BD FACSaria™ II cell sorter (BD Biosciences).

**In vivo xenograft experiments and animal care**

MCF-7 cells were transfected with indicated shRNAs (5 × 10⁴, 5 × 10⁵, 5 × 10⁶ cells, n = 8/group), suspended in PBS, and mixed with Matrigel (1:1, BD Biosciences, catalog no. 354234). Tumor volume was measured every 3 days from the time tumor palpable until the animals were sacrificed (week 6). Palpable tumor volumes and tumor volumes at the end point (week 6) were measured by width and length with a Vernier caliper and calculated by formula Volume = (Length × Width × Width)/2. SCID mice (BALB/cJNju-Foxn1nu/Nju) were purchased from Nanjing Biomedical Research Institute of Nanjing University. All animal studies were carried out in accordance with the guidelines for the care and use of laboratory animals of Southeast University.

**Data mining**

The data sets of breast cancer patients were downloaded from the publicly available GEO databases using the BioProject ID PRJNA376644. The data were analyzed starting with the raw FASTQ files. After quality checking, alignment was performed with TopHat, and the reference was GRCh38 (hg38). The quantification of the genes was carried out with HTSeq-count, and then the differential expression was called with DESeq.
References

1. Clevens, H. (2011) The cancer stem cell: premises, promises and challenges. *Nat. Med.* 17, 313–319

2. Boumahdi, S., Driessens, G., Lapouge, G., Rorive, S., Nassar, D., Le Mercier, M., Delatte, B., Cauwe, A., Lengelé, S., Nikusi, E., Brohée, S., Salmon, I., Dubois, C., del Marmol, V., Fukès, F., et al. (2014) SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature 511*, 246–250

3. Santini, R., Pietrobono, S., Pandolfi, S., Montagnani, V., D’Amico, M., Penachioni, J. Y., Vinci, M. C., Borgognoni, L., and Stecca, B. (2014) SOX2 regulates self-renewal and tumorigenicity of human melanoma-initiating cells. *Oncogene 33*, 4697–4708

4. Annovazzi, L., Mellai, M., Caldera, V., Valente, G., and Schiffer, D. (2011) SOX2 expression and amplification in gliomas and glioma cell lines. *Cancer Genomics Proteomics 8*, 139–147

5. Leis, O., Eguiara, A., Lopez-Arribillaga, E., Alberdi, M. J., Hernandez-Garcia, S., Elorriaga, K., Pandiella, A., Rezola, R., and Martin, A. G. (2012) Sox2 expression in breast tumours and activation in breast cancer stem cells. *Oncogene 31*, 1354–1365

6. Piva, M., Domenici, G., Iriondo, O., Rabano, M., Simões, B. M., Comaillas, V., Barredo, I., López-Ruiz, J. A., Zabalza, I., Kypta, R., and Vivanco, M. (2014) Sox2 promotes tamoxifen resistance in breast cancer cells. *EMBO Mol. Med. 6*, 66–79

7. Zhang, C., Ye, X., Zhang, H., Ding, M., and Deng, H. (2007) GATA factors induce mouse embryonic stem cell differentiation toward extraembryonic endoderm. *Stem Cells Dev. 16*, 605–613

8. Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J., and Niwa, H. (2002) Differentiation of embryonic stem cells is induced by GATA factors. *Gene Dev. 16*, 784–789

9. Turbendian, H. K., Gordillo, M., Lu, J., Kang, G., Liu, T. C., Thiery, J. P., and Chen, L. (2014) A central role for TRPS1 in the control of follicle morphogenesis. *Development 140*, 1639–1644

10. Zheng, R., and Blobel, G. A. (2010) GATA transcription factors and cancer. *Genes Cancer 1*, 1178–1188

11. Lentjes, M. H., Nissen, H. E., Akiyama, H., de Bruïne, A. P., Melotte, V., and van Engeland, M. (2016) The emerging role of GATA transcription factors in development and disease. *Expert Rev. Mol. Med. 18*, e3

12. Itoh, S., Kanno, S., Gai, Z., Suemoto, H., Kawakatsu, M., Tanishima, H., Rogers, K., Ward, J. M., Mori, S., Adams, D. J., Jenkins, N. A., Copeland, N. G., et al. (2017) Transposon insertional mutagenesis in mice identifies human breast cancer susceptibility genes and signatures for stratification. *Proc. Natl. Acad. Sci. U.S.A. 114*, E2215–E2224

13. Liu, K., Jin, B., Zhao, M., Yang, X., Chen, M., Gao, A., Liu, F., Que, J., and Lan, X. (2013) The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. *Cell. Signal. 25*, 1264–1271

14. Chen, J. Q., Bao, Y., Litton, J., Xiao, L., Zhang, H. Z., Warneke, C. L., Wu, Y., Shen, X., Wu, S., Katz, R. L., Sahin, A., Bondy, M., Murray, J. L., and Radvanyi, L. (2011) Expression and relevance of TRPS-1: a new GATA transcription factor in breast cancer. *Hormones Cancer 2*, 132–143

15. Charafe-Jauffret, E., Ginestier, C., Iovino, F., Wicinska, J., Cervera, N., Finietti, P., Hur, M. H., Diebel, M. E., Monville, F., Dutcher, J., Brown, M., Viens, P., Xerri, L., Bertucci, F., Stassi, G., et al. (2009) Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Res. 69*, 1302–1313

16. Rangel, R., Lee, S. C., Hon-Kim Ban, K., Guzman-Rojas, L., Mann, M. B., Newberg, J. Y., Kodama, T., McNoe, L. A., Selvanesan, L., Ward, J. M., Rust, A. G., Chin, K. Y., Black, M. A., Jenkins, N. A., and Copeland, N. G. (2016) Transposon mutagenesis identifies genes that cooperate with mutant Pten in breast cancer progression. *Proc. Natl. Acad. Sci. U.S.A. 113*, E7749–E7758

17. Chen, L., Jenjaroenpun, P., Pillai, A. M., Ivshina, A. V., Ow, G. S., Efthimiou, M., Zhiqun, T., Tan, T. Z., Lee, S. C., Rogers, K., Ward, J. M., Mori, S., Adams, D. J., Jenkins, N. A., Copeland, N. G., et al. (2017) Transposon insertional mutagenesis in mice identifies human breast cancer susceptibility genes and signatures for stratification. *Proc. Natl. Acad. Sci. U.S.A. 114*, E2215–E2224

18. Malik, T. H., Shoichet, S. A., Latham, P., Kroll, T. G., Peters, L. L., and Shvidrasani, R. A. (2001) Transcriptional repression and developmental functions of the atypical vertebrate GATA protein TRPS1. *EMBO J. 20*, 1715–1725

19. Rodriguez-Pinilla, S. M., Sarrio, D., Moreno-Bueno, G., Rodriguez-Gil, Y., Martinez, M. A., Hernandez, L., Hardisson, D., Reis-Filho, J. S., and Palacios, J. (2007) Sox2: a possible driver of the basal-like phenotype in sporadic breast cancer. *Mod. Pathol. 20*, 474–481

20. Bussolati, G., and van Engeland, M. (2016) The emerging role of GATA transcription factors in development and disease. *Nat. Rev. Mol. Med. 17*, 355–363

21. Papapetrou, E. P., Tomishima, M. J., Chambers, S. M., Mica, Y., Reed, E., Menon, J., Tabar, V., Mo, Q., Studer, L., and Sadelain, M. (2009) Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc. Natl. Acad. Sci. U.S.A. 106*, 12759–12764

22. Li, Q., Housman, D., Robertson, K. H., Phillips, S. H., Huang, L., Xia, Y., Shen, X., Wu, S., Katz, R. L., Sahin, A., Bondy, M., Murray, J. L., et al. (2009) Generation of induced pluripotent stem cells from human cord blood. *Cell Stem Cell 5*, 434–441

23. Dong, S., Shi, X., Zhang, H., Ding, M., and Deng, H. (2007) GATA factors induce mouse embryonic stem cell differentiation toward extraembryonic endoderm. *Stem Cells Dev. 16*, 605–613

24. Basu-Roy, U., Seo, E., Ramanathanuram, L., Rapp, T. B., Perry, J. A., Orkin, S. H., Mansukhani, A., and Basilico, C. (2012) Sox2 maintains self-renewal of tumor-initiating cells in osteosarcomas. *Oncogene 31*, 2270–2282

25. Haase, A., Olmer, R., Schwarke, K., Wunderlich, S., Merkert, S., Hess, C., Zeidler, D., and Zeidler, D. (2014) SOX2: a possible driver of the basal-like phenotype in sporadic breast cancer. *Stem Cells Transl. Med. 3*, 511–519

26. Papapetrou, E. P., Tomishima, M. J., Chambers, S. M., Mica, Y., Reed, E., Menon, J., Tabar, V., Mo, Q., Studer, L., and Sadelain, M. (2009) Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc. Natl. Acad. Sci. U.S.A. 106*, 12759–12764

27. Zhao, H. X., Li, Y., Jin, H. F., Xie, L., Liu, C., Jiang, F., Luo, Y. N., Yin, G. W., Li, Y., Wang, J., Li, L. S., Yao, Y. Q., and Wang, X. H. (2010) Rapid and efficient reprogramming of human amnion-derived cells into pluripotency by three factors OCT4/SOX2/NANOG. *Differentiation 80*, 123–129

28. Donou, G., Al-Haji, M., Abdallah, W. M., Clarke, M. F., and Wicha, M. S. (2003) Stem cells in normal breast development and breast cancer. *Cell Proliferation 36*, 59–72

29. Shaw, F. L., Harrison, H., Spence, K., Ablett, M. P., Simões, B. M., Farnie, G., et al. (2013) Sox2 promotes tamoxifen resistance in breast cancer cells. *Mol. Med. 19*, 1354–1365

30. Knowles, D. M., Williams, C. A., and Evan, G. (2004) Identification of human brain tumour initiating cells. *Nature 432*, 396–401
31. Reya, T., Morrison, S. J., Clarke, M. F., and Weissman, I. L. (2001) Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 CrossRef Medline

32. Vidal, S. J., Rodriguez-Bravo, V., Galsky, M., Cordon-Cardo, C., and Domingo-Domenech, J. (2014) Targeting cancer stem cells to suppress acquired chemotherapy resistance. *Oncogene* **33**, 4451–4463 CrossRef Medline

33. Liu, S., Cong, Y., Wang, D., Sun, Y., Deng, L., Liu, Y., Martin-Trevino, R., Shang, L., McDermott, S. P., Landis, M. D., Hong, S., Adams, A., D’Angelo, R., Ginestier, C., Charafe-Jauffret, E., et al. (2014) Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports* **2**, 78–91 CrossRef Medline

34. Savinainen, K. J., Linja, M. J., Saramäki, O. R., Tammela, T. L., Chang, G. T., Brinkmann, A. O., and Visakorpi, T. (2004) Expression and copy number analysis of TRPS1, EIF3S3 and MYC genes in breast and prostate cancer. Br. J. Cancer **90**, 1041–1046 CrossRef Medline

35. Wang, Y. Z., Zhang, J., Wu, L. L., Liu, W. G., Wei, G. Y., Gong, X., Liu, Y., Ma, Z. F., Ma, F., Thiery, J. P., and Chen, L. M. (2018) Tricho-rhino-phalangeal syndrome 1 protein functions as a scaffold required for ubiquitin-specific protease 4-directed histone deacetylase 2 de-ubiquitination and tumor growth. *Breast Cancer Res.* **20**

36. Kouros-Mehr, H., Slorach, E. M., Sternlicht, M. D., and Werb, Z. (2006) GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* **127**, 1041–1055 CrossRef Medline

37. Usary, J., Llaca, V., Karaca, G., Presswala, S., Karaca, M., He, X., Langerod, A., Kåresen, R., Oh, D. S., Dressler, L. G., Lenning, P. E., Strausberg, R. L., Chanoick, S., Berresen-Dale, A. L., and Perou, C. M. (2004) Mutation of GATA3 in human breast tumors. *Oncogene* **23**, 7669–7678 CrossRef Medline

38. Yu, S., Cai, X., Wu, C., Liu, Y., Zhang, J., Gong, X., Wang, X., Wu, X., Zhu, T., Mo, L., Gu, J., Yu, Z., Chen, J., Thiery, J. P., Chai, R., et al. (2017) Targeting HSP90-HDAC6 regulating network implicates precision treatment of breast cancer. *Int. J. Biol. Sci.* **13**, 505–517 CrossRef Medline

39. Lee, T. I., Johnstone, S. E., and Young, R. A. (2006) Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat. Protocols* **1**, 729–748 CrossRef Medline

40. Hu, Y., and Smyth, G. K. (2009) ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347**, 70–78 CrossRef Medline

TRPS1 suppresses cancer stem cell functions