OPEN

Oncogene (2014) **33,** 4273–4278

www.nature.com/onc

SHORT COMMUNICATION

An *in vivo* RNAi screen identifies *SALL1* as a tumor suppressor in human breast cancer with a role in *CDH1* regulation

J Wolf[1], K Müller-Decker[2], C Flechtenmacher[3], F Zhang[4], M Shahmoradgoli[4], GB Mills[4], JD Hoheisel[1] and M Boettcher[1]

The gold standard for determining the tumorigenic potential of human cancer cells is a xenotransplantation into immunodeficient mice. Higher tumorigenicity of cells is associated with earlier tumor onset. Here, we used xenotransplantation to assess the tumorigenic potential of human breast cancer cells following RNA interference-mediated inhibition of over 5000 genes. We identify 16 candidate tumor suppressors, one of which is the zinc-finger transcription factor *SALL1*. Analyzing this particular molecule in more detail, we show that inhibition of *SALL1* correlates with reduced levels of *CDH1*, an important contributor to epithelial-to-mesenchymal transition. Furthermore, *SALL1* expression led to an increased migration and more than twice as many cells expressing a cancer stem cell signature. Also, *SALL1* expression correlates with the survival of breast cancer patients. These findings cast new light on a gene that has previously been described to be relevant during embryogenesis, but not carcinogenesis.

Oncogene (2014) **33,** 4273–4278; doi:10.1038/onc.2013.515; published online 2 December 2013

Keywords: *In vivo* RNAi screening; shRNA screen; tumor suppressor; *SALL1*; *CDH1*; EMT

INTRODUCTION

Currently, the best assay to determine the tumorigenic potential of tumor cells involves xenotransplantation of different subpopulations of cancer cells into highly immunosuppressive animals.[1,2] This approach has been used successfully for a functional *in vivo* identification of tumor suppressors in several tumor types.[3,4] Taking advantage of a pooled, lentiviral short hairpin RNA (shRNA) library, we performed a large-scale *in vivo* RNA interference (RNAi) screen in mice to find genes that act as tumor suppressors in human breast cancer cells. As a part of this study, *SALL1* was identified, which is one of the four human family members of the spalt family. Members of the spalt family are highly conserved zinc-finger transcription factors, present from *Caenorhabditis elegans* to vertebrates, with regulatory functions in organogenesis, limb formation and cell-fate assignment during neural development.[5] Mutations in the human *SALL1* gene have been associated with Townes–Brocks syndrome, a rare, dominantly inherited disorder, characterized by limb, ear, anal and renal abnormalities.[6] In mice, Sall1 acts as a regulator of canonical Wnt signaling during kidney development.[7] Furthermore, it has been found that, in mouse embryonic stem cells, Sall1 physically interacts with Nanog and Sox2 and it has further been suggested to be a novel component of stemness.[8] Although the importance of Sall1 during mouse embryogenesis has long been recognized, the first data suggesting the involvement of its human ortholog in carcinogenesis have only been published recently. It has been shown that *SALL1* silencing via promoter hypermethylation is associated with human breast cancer.[9] A functional role of *SALL1* in breast cancer, however, had not been described before. Interestingly, our data show that *SALL1* expression correlates with the expression of the gene *CDH1*, a key factor during epithelial-to-mesenchymal transition (EMT).[10] EMT is a tightly controlled developmental mechanism involved in processes including embryogenesis, tissue repair and wound healing.[11] Its induction leads to the loss of polarized epithelial and the acquisition of mesenchymal motile cell phenotypes.[11] During carcinogenesis, EMT facilitates cancer cell migration and metastasis.[12] Moreover, induction of EMT in cancer cells promotes their tumorigenicity,[13] and cancer cells that have undergone EMT gain cancer stem cell properties.[14] The finding that *SALL1* expression is correlative with the expression of *CDH1* is consistent with its tumor suppressive function and suggests its potential involvement in EMT.

RESULTS AND DISCUSSION

The triple negative human breast cancer cell line SUM-149 was transduced with the pooled, lentiviral Decipher library module 1.[15] Transduction conditions were adjusted to ensure a maximum of one integration event per target cell. The library used consists of 27 494 shRNA expression constructs, targeting 5045 genes for knockdown by five to six dissimilar shRNA sequences each. Upon infection, each construct integrates into the genome of the host cells, expressing the relevant small interfering RNA continuously. For a pooled RNAi screen with all constructs, the number of cells that express a particular shRNA is critical to ensure statistical significance. As not every cancer cell has the potential to form a tumor, we conducted a preliminary experiment to determine the fraction of tumorigenic SUM-149 cells. For this purpose, 6000 cells each were injected subcutaneously into eight NOD SCID mice. It was found that in six out of eight cases, tumors developed within 25 weeks post injection. From this, it could be concluded that 48 000 cells harbored a minimum of 6 cells capable of initiating a tumor. On the basis of this estimation, a total of 10^9 cells were transduced with the Decipher library. As a result, 36 000 cells, of which at least 4 cells should have a tumor initiating capacity,

[1]Division of Functional Genome Analysis, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany; [2]Core Facility Tumor Models, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany; [3]Institute of Pathology, Heidelberg University Hospital, Heidelberg, Germany and [4]Department of Systems Biology, MD Anderson Cancer Centre, Houston, TX, USA. Correspondence: Dr JD Hoheisel, Division of Functional Genome Analysis, Deutsches Krebsforschungszentrum (DKFZ), Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany.
E-mail: j.hoheisel@dkfz.de
Received 18 March 2013; revised 10 October 2013; accepted 21 October 2013; published online 2 December 2013

expressed each shRNA in the library. Although this figure is too low to allow negative selection screening, it does ensure that each shRNA is represented sufficiently in the pool of tumor cells to identify shRNAs that increase their tumorigenic potential. Following transduction, 10^7 cells each were injected subcutaneously into the flanks of 50 NOD SCID mice. The scheme shown in Figure 1a illustrates the overall workflow. Tumor development to a diameter of 1.5 cm took on average 57 ± 12 days with a take rate of 0.86. Tumors were divided into two groups and analyzed separately via next-generation sequencing of barcode sequences, which are unique for each particular shRNA expression construct in the library. Consequently, barcode sequences could be used to identify each shRNA expression construct and thus the number of cells containing it. Barcode read counts from tumors were divided by baseline read counts from cells 3 days post transduction. Finally, *z*-scores—the distance of the value from the mean in s.d. units—were computed for each shRNA expression construct (Supplementary Table S1). Correlation between biological replicates was found to be $r = 0.66$. The *z*-scores of all shRNA expression constructs in the library are shown in Figure 1b.

The Decipher library contains 21 negative control shRNAs targeting the expression of the gene luciferase. The twofold s.d. of

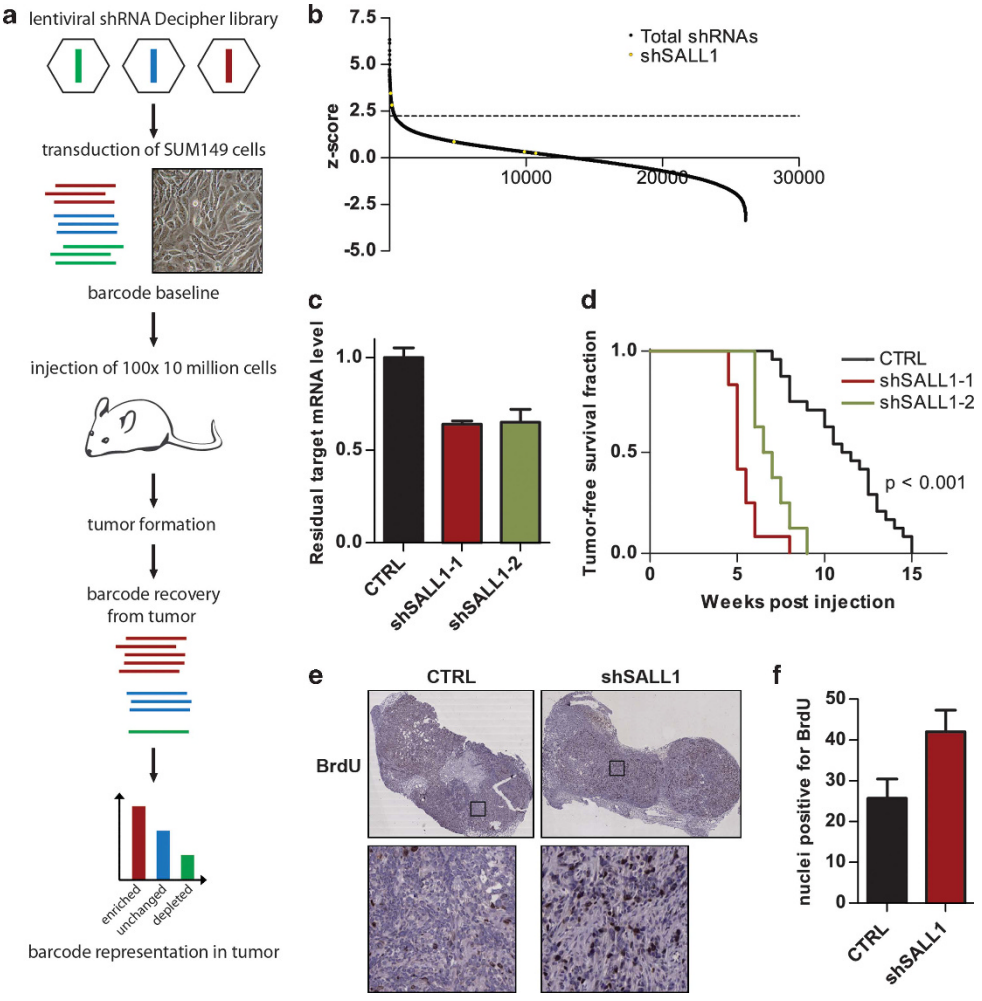

Figure 1. Barcoded *in vivo* RNAi screen with shRNA constructs. (**a**) Scheme of the process: cells were transduced with Decipher library module 1 (Cellecta Inc., Mountain View, CA, USA). Into each flank of 50 NOD SCID mice, 10^7 stably transduced SUM-149 cells suspended in 200 µl PBS/Matrigel (BD Biosciences, Billerica, MA, USA) (1:1, v/v) were injected subcutaneously. Following tumor formation, tumors were homogenized and genomic DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Amplification of barcode sequences was achieved by two rounds of PCR according to a protocol described in detail elsewhere.[35] The amplified sequences were purified by means of PCR purification and gel extraction kits (#28104 and #28704, Qiagen). Then, barcode representation was quantified using next-generation sequencing on GAIIx machines (Illumina Inc., San Diego, CA, USA). Barcodes associated with shRNAs that promote tumor formation became enriched (illustrated in red). (**b**) Ranking of the *z*-scores from each barcode in the library. The dashed line indicates the cutoff defined by the twofold s.d. of the average *z*-score of the 21 negative control shRNAs. The *z*-scores from five shRNA expression constructs targeting *SALL1* are shown in orange. Genes represented by at least two shRNAs with *z*-scores above the cutoff were considered candidate genes and are listed in Table 1. (**c**) Residual mRNA levels following 5 days of expression of the two shRNAs targeting *SALL1* in SUM-149 cells. Reverse transcription of RNA and PCR was performed in one step using the Quantifast SYBR Green RT-PCR Kit (Qiagen) on a LightCycler 480 system (Roche, Basel, Switzerland). The QuantiTect Primer Assay (Qiagen) was used for specific target gene amplification. (**d**) Tumor-free survival of NOD SCID gamma mice following orthotopic injection of 40 000 SUM-149 cells with reduced *SALL1* expression (sh*SALL1*-1/ -2) relative to control (shCTRL). Significance values from Kaplan–Meier plots were calculated by means of the Wilcoxon test, using GraphPad Prism software (Graphpad, La Jolla, CA, USA). (**e**) Anti-bromodeoxyuridine (BrdU) antibody staining from tumors with inhibited *SALL1* (sh*SALL1*) or control tumor (shCTRL), respectively. Black boxes indicate enlarged areas shown below. For BrdU staining, mice were injected with BrdU 2 h before tumor collection, and samples were prepared according to the protocol from the Cell Proliferation Kit (G&E Healthcare, Wauwatosa, WI, USA). (**f**) Quantification of six stainings such as shown in **e**.

the *z*-scores of these shRNAs was used to define the cutoff for candidate shRNAs. The cutoff was 2.24 and is indicated by the dashed line in Figure 1b. To reduce the risk of off-target effects, genes were considered candidates only when at least two targeting shRNAs showed a *z*-score above this value. In total, 16 genes were found to meet these criteria (Table 1). To exclude that these candidates drove tumor growth only because of an increased proliferation *in vitro*, we cultured cells transduced with the shRNA library for 14 days and determined the construct pool representation before and after the culture period. By means of comparison, the proliferative nature of each gene was determined. Knockdown of 12 candidate genes had no significant impact on proliferation, whereas the inhibition of the genes *ANP32E*, *MUTYH*, *GDF6* and *RAD54L* even led to reduced proliferation (Table 1).

Two of the shRNA molecules identified in this screen targeted the expression of *SALL1*. This gene was chosen for subsequent analyses because it belongs to the group of zinc-finger transcription factors[16] and has been implicated in breast cancer recently.[9] The two shRNAs were sub-cloned individually into the expression vector pRSI9, yielding the constructs shSALL1-1 and shSALL1-2. SUM-149 cells were transduced with each expression construct. The residual target messenger RNA (mRNA) levels were found to be reduced 5 days post transduction (Figure 1c). Cells were then injected into the mammary fat pad of NOD SCID gamma mice (40 000 cells per animal). The time between injection and tumor onset was recorded. The inhibition of *SALL1* by either shRNA resulted in significantly decreased tumor-free survival periods (Figure 1d). Measurement of tumor growth kinetics showed clear differences. Although these results further support a tumor suppressive role for *SALL1*, they were not significant. This can be explained by the large s.d. resulting from the variation in time to tumor onset as well as the biological variation between individual animals. Other shRNA sequences targeting the candidate genes *EMR3* and *GPRC5D* were validated in the same way and yielded similar results (Supplementary Figure S1), indicating the accuracy of the initial screen results. Next, tumors that had formed from cells with inhibited *SALL1* expression were compared with control tumors via immunohistochemistry. Figure 1e shows sections from representative tumors chased with bromodeoxyuridine. Tumors with inhibited *SALL1* expression is more frequently stained positive for bromodeoxyuridine (Figure 1f), indicating a larger fraction of mitotically active cells.

Table 1. Candidate tumor suppressor genes identified by means of *in vivo* RNAi screen

Gene symbol	*Description*	In vitro *proliferation*	
		log_2-fold change	P-*value*
AMPD1	Adenosine monophosphate deaminase 1	0.008	0.97
ANP32E	Acidic (leucine-rich) nuclear phosphoprotein 32 family	−0.267	0.003
CAPN1	Calpain 1, (mu/I) large subunit	−0.203	0.19
CASP8	Caspase 8, apoptosis-related cysteine peptidase	−0.270	0.07
DPM1	Dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic subunit	0.018	0.97
ELK1	ELK1, member of ETS family	0.026	0.61
EMR3	EGF-like module containing, mucin-like, hormone receptor-like 3	−0.124	0.34
GDF6	Growth differentiation factor 6	−0.234	0.004
GPRC5D	G-protein-coupled receptor, family C, group 5, member D	0.112	0.36
GRB14	Growth factor receptor-bound protein 14	0.148	0.29
MUTYH	mutY homolog (*Escherichia coli*)	−0.195	0.019
PMM1	Phosphomannomutase 1	−0.088	0.40
RAD54L	RAD54-like (*Saccharomyces cerevisiae*)	−0.409	0.005
RARG	Retinoic acid receptor gamma	−0.002	0.84
SALL1	Sal-like 1 (*Drosophila*)	0.212	0.07
SLC1A4	Solute carrier family 1 (glutamate/neutral amino acid transporter), member 4	−0.115	0.166

Listed values indicate inhibited (negative) or increased (positive) proliferation with associated *P*-values over a period of 14 days as determined by an additional cell survival screen in SUM-149 cells.

SALL1 is known to act as a transcriptional repressor in non-cancer cells.[16] To identify genes whose expression levels were changed following *SALL1* inhibition, a microarray expression profile was performed. SUM-149 cells with inhibited *SALL1* expression that was induced by the constructs shSALL1-1 or shSALL1-2, respectively, were compared to control cells (Supplementary Table S2). In total, 200 genes were identified to be significantly ($P<10^{-10}$) up- or downregulated following the expression of either shRNA. These genes, together with the observed expression-changes caused by both shRNAs, are summarized in Supplementary Table S3. Interestingly, the inhibition of *SALL1* was found to occur jointly with altered expression of several important factors involved in EMT, including *CDH1* (Thiery *et al.*[17]), *CDH2* (Kalluri and Weinberg[18]), *VIM*[18] and *MSN*.[19,20] Relevant variations found during the microarray analysis were validated by means of quantitative RT–PCR (Figure 2a). In SUM-149 cells, expression of each shRNA targeting *SALL1* led to reduced *SALL1* and *CDH1* expression and increased *CDH2*, *VIM* and *MSN* expression, which is a typical expression signature for mesenchymal cells. In addition, *SALL1* inhibition was confirmed to correlate with increased expression of the two oncogenic CCN family members *CTGF* and *CYR61*, both of which are known to have a critical role in breast carcinogenesis.[21,22] Furthermore, mRNA levels of two putative tumor suppressor genes, retinoic acid receptor responders *RARRES1* (Jing *et al.*[23]) and *RARRES3* (Hsu *et al.*[24]), were confirmed to be reduced following *SALL1* inhibition. Moreover, we analyzed protein levels of two EMT markers, E-cadherin and vimentin, and found reduced E-cadherin and increased vimentin levels following *SALL1* knockdown (Figure 2b or c).

To investigate the effects of *SALL1* inhibition in different genetic backgrounds, *SALL1* expression (Supplementary Figure S1c) and residual target mRNA levels following *SALL1* inhibition were determined via quantitative RT–PCR in the basal breast carcinoma cell lines SUM-159 and MDA-MB-231, as well as in luminal MCF-7 and SKBR3 cells. Similar to SUM-149 cells, the inhibition of *SALL1* by constructs shSALL1-1 or shSALL1-2 led to reduced *CDH1* levels in all investigated breast cancer cell lines (Figure 2d). Given the diverse genetic background of these cell lines, this effect of *SALL1* inhibition appears to be independent of genomic aberrations.

Furthermore, the impact of *SALL1* inhibition on the migratory phenotype of SUM-149 was investigated. To this end, wound-healing experiments using cells transduced with the constructs shSALL1-1, shSALL1-2 and a control were performed. As shown in Figure 3a, inhibition of *SALL1* increased the migratory potential of SUM-149.

Mani *et al.*[14] have previously demonstrated a link between EMT and an increase in breast cancer cells with stem-like properties such as $CD44^+/CD24^{-/low}$. Moreover, Gupta *et al.*[25] identified a distinct sub-population of SUM-149 cells with cancer stem cell-like properties displaying the surface molecule signature $CD44^+/CD24^{-/low}/EpCAM^{-/low}$. To determine whether inhibition of

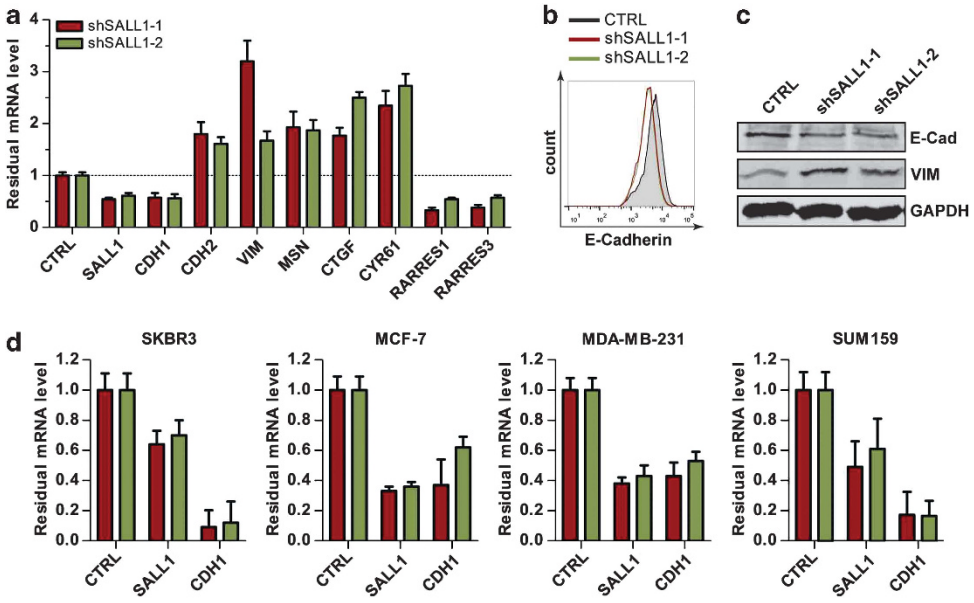

Figure 2. *SALL1* expression correlates with *CDH1* levels in breast cancer cell lines. (**a**) Fold-changes of mRNA levels of indicated genes following 5 days of expression of two shRNAs targeting *SALL1* in SUM-149 cells relative to control cells. (**b**) SUM-149 cells transduced with shSALL1-1, shSALL1-2 or a control construct were stained with E-cadherin-PE (Miltenyi Biotec, Bergisch Gladbach, Germany) and analyzed using flow cytometry. (**c**) Protein levels were determined by means of western blotting. Membranes were probed with antibodies against E-cadherin (Abgent, San Diego, CA, USA), vimentin (Cell Signaling, Danvers, MA, USA) and GAPDH (Abcam, Cambridge, UK), detected using peroxidase-conjugated antibodies (Sigma-Aldrich, St Louis, MO, USA) and ECL (Thermo, Waltham, MA, USA). (**d**) Reduction of the mRNA levels of *CDH1* and *SALL1* following 5 days of expression of the two shRNAs targeting *SALL1* in the indicated breast cancer cell lines SKBR3, MCF-7, MDA-MB-231 and SUM-159. All measurements were relative to control cells with a non-inhibiting shRNA construct.

SALL1 influences the percentage of cells expressing a cancer stem cell signature, CD44/CD24/EpCAM expression levels were compared between knockdown and control cells using flow cytometry.[26] SUM-149 cells transduced with shSALL1-1 and shSALL1-2 exhibited more than twice as many $CD44^+/CD24^{-/low}/EpCAM^{-/low}$ cells in comparison to the control population (Figure 3b).

Analysis of *SALL1* expression in breast cancer patient samples was consistent with a tumor suppressive role of *SALL1*. The analysis of two independent patient datasets[27,28] revealed that high *SALL1* mRNA levels were associated with a significantly increased relapse-free survival ($P = 0.00048$), overall survival ($P = 0.0027$), metastasis-free survival ($P = 0.0071$) as well as tumor-free survival ($P = 0.011$) (Figure 4a or b).

On the basis of gene expression profiles, breast cancer can be grouped into several distinct subtypes.[29] The basal-like subtype contains an increased percentage of $CD44^+/CD24^-$ and $ALDH1^+$ cancer stem cells.[30] A subtype-specific analysis of *SALL1* expression levels revealed that mRNA levels were significantly lower ($P < 2 \times 10^{-9}$) in the most aggressive basal-like subtype[29] when compared with other breast cancer subtypes (Figure 4c), again consistent with the association with patient outcome.

Understanding the molecular mechanisms that drive carcinogenesis is essential to understand the development of human tumors. Tumor suppressor genes are known to have an important role in this process because their loss of function enhances the tumorigenic potential of cancer cells. Here, we conducted a large-scale *in vivo* RNAi screen aimed at the identification of novel breast tumor suppressor genes. Our results attribute a tumor suppressive role to the transcriptional repressor *SALL1* in the background of the triple negative, tumorigenic breast cancer cell line SUM-149. Although in this study, the effects of *SALL1* inhibition on non-tumorigenic cells, such as HMLER, were not determined, it was clearly shown that *SALL1* inhibition leads to a more aggressive phenotype in SUM-149 cells.

SALL1 has an important role during embryonic kidney development and affects Wnt/beta-catenin signaling,[7] a pathway commonly activated during EMT.[18] Furthermore, other studies have demonstrated methylation of the *SALL1* promoter in breast tumors, colorectal cancer, non-small-cell lung carcinoma and acute lymphocytic leukemia.[9] Also, it has been shown that the region 16q12.1, in which *SALL1* is located, often gets deleted in breast tumors.[31,32] Both observations, inactivation through methylation and loss of heterozygosity of *SALL1* in different cancers, are supportive of its tumor suppressive function.

In line with these findings, we found that *SALL1* expression is associated with the expression levels of a number of genes involved in EMT, namely *CDH1*, *CDH2*, *VIM* and *MSN*. Furthermore, *SALL1* inhibition led to reduced expression of *CDH1* in four additional breast cancer cell lines. It has previously been found that inhibition of *CDH1* is sufficient to induce human epithelial breast cells to undergo EMT,[33,34] emphasizing the crucial role *CDH1* has during this process. Moreover, inhibition of *SALL1* led to reduced protein levels of E-cadherin, encoded by *CDH1*, whereas vimentin levels increased; both changes are typical for EMT.[33] In this respect, it is important to note that *SALL1* has been shown to inhibit the induction of Goosecoid, a repressor of *CDH1* (Thiery *et al.*[17]), during embryoid body differentiation.[8] This provides one possible route by which *SALL1* might regulate the expression of *CDH1*. Furthermore, it is known that induction of EMT leads to an increase of cells with a $CD44^+/CD24^-$ phenotype.[14] After *SALL1* knockdown, we detected a significant increase in the fraction of $CD44^+/CD24^-$.

It is important to mention that the data presented here does not provide evidence for a causative role of *SALL1* in EMT. It does, however, clearly show that there is a correlative link between the expression of *SALL1* and that of *CDH1* in five different human breast cancer cell lines. Also, increased *in vitro* invasiveness and expression of a cancer stem cell marker signature following *SALL1* inhibition are phenotypes frequently associated with EMT,

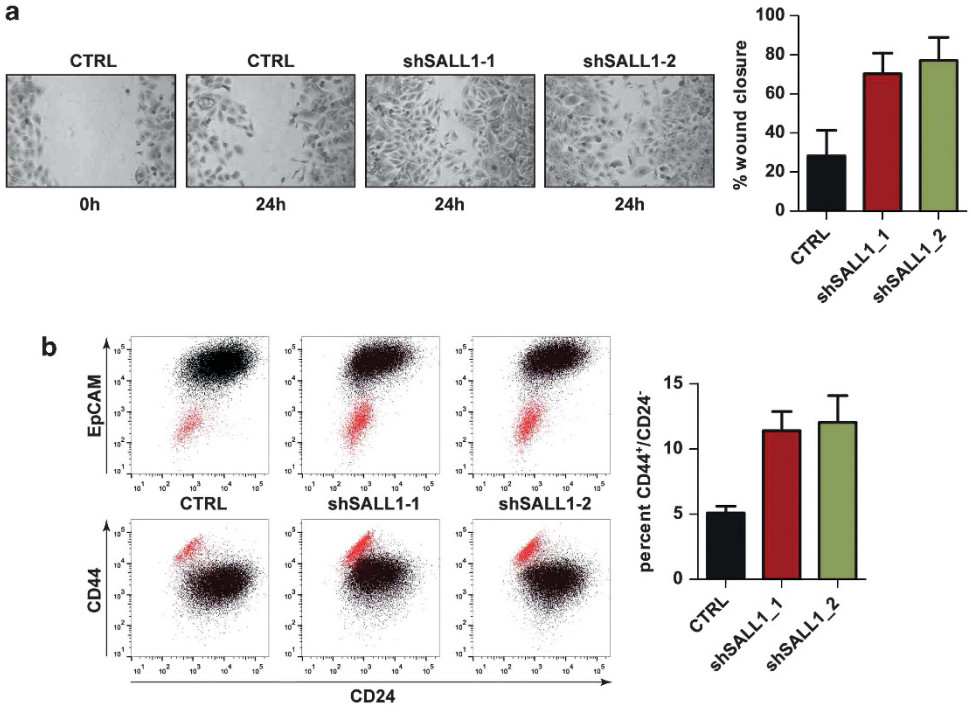

Figure 3. Impact of *SALL1* knockdown on cell migration and CSC population. (**a**) Effects of *SALL1* inhibition on the cell migration of SUM-149 cells. Transduced cells were scratched to create a gap of 700 μm and washed twice with PBS. Directly after creation of the gap and again after 24 h, pictures were taken, and the gap size was analyzed using ImageJ software (http://rsb.info.nih.gov/ij/). Quantifications from 24 such wound-healing assays are shown on the right. (**b**) Flow cytometry analysis of CD44, CD24 and EpCAM expression of cells, whose *SALL1* expression was modified by transduction with shSALL1-1 and shSALL1-2 or a control construct (CTRL). SUM-149 cells were stained with CD44-PE/Cy7, CD24-FITC, EpCAM-APC (Becton Dickinson, Franklin Lakes, NJ, USA).

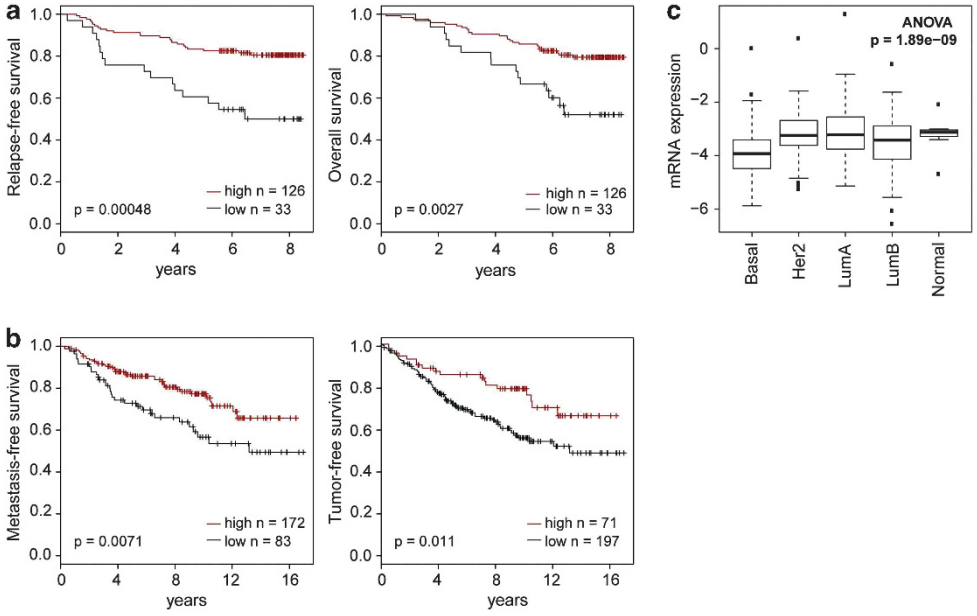

Figure 4. Association of *SALL1* expression with patient survival and breast cancer subtypes. Relapse-free and overall survival, as well as the metastasis-free and tumor-free fractions of breast cancer patients with high or low *SALL1* expression levels are shown. A Cox proportional hazard regression model was used for a univariate survival analysis of the gene expression data sets of (**a**) Pawitan *et al.*[27] and (**b**) Loi *et al.*[28] The samples were sorted according to the expression of *SALL1*, then all cutoffs in the central 60% of samples were assessed for correlation with outcomes. The split with the lowest *P*-value was chosen. (**c**) The Cancer Genome Atlas (TCGA) data set was used to analyze the mRNA expression levels of *SALL1* across breast cancer tumor subtypes (downloaded in August 2012 at https://tcga-data.nci.nih.gov/tcga/). One-way ANOVA was used to compare the level of *SALL1* mRNA expression among the different breast cancer subtypes (luminal (LumA and LumB), basal and HER2) and normal-like tissue.

although not exclusively. Their occurrence might thus be correlative with reduced *SALL1* levels, but have no impact on the tumorigenic potential of cells or EMT. Hence, whether or not *SALL1* is actually involved in the regulation of the highly complex process of EMT is yet to be firmly demonstrated.

In summary, we performed a large-scale *in vivo* RNAi screen, which led to the identification of *SALL1* as a novel tumor suppressor gene in human breast cancer cells. We show that *SALL1* expression is associated with the expression of a central regulator of EMT, namely *CDH1*, and that *SALL1* expression correlates with the survival of breast cancer patients. These findings depict *SALL1* as an exciting new tumor suppressor warranting closer investigation, especially regarding its potential involvement in EMT, as well as in breast cancer development.

ANIMAL STUDIES

The animal studies were approved by the local ethics committee at the Regierungspräsidium Karlsruhe (G74/11, G244/11).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by a grant from the DKFZ intramural funding program awarded to MB, and a DKFZ PhD scholarship awarded to JW. Lentiviral shRNA libraries were kindly provided and developed by Cellecta based on NIH-funded research grant support 44RR024095, 44HG003355.

REFERENCES

1 Quintana E, Shackleton M, Sabel MS, Fullen DR, Johnson TM, Morrison SJ. Efficient tumour formation by single human melanoma cells. *Nature* 2008; **456**: 593–598.
2 Krelin Y, Berkovich L, Amit M, Gil Z. Association between tumorigenic potential and the fate of cancer cells in a syngeneic melanoma model. *PLoS One* 2013; **8**: e62124.
3 Zender L, Xue W, Zuber J, Semighini CP, Krasnitz A, Ma B *et al.* An oncogenomics-based *in vivo* RNAi screen identifies tumor suppressors in liver cancer. *Cell* 2008; **135**: 852–864.
4 Bric A, Miething C, Bialucha CU, Scuoppo C, Zender L, Krasnitz A *et al.* Functional identification of tumor-suppressor genes through an in vivo RNA interference screen in a mouse lymphoma model. *Cancer Cell* 2009; **16**: 324–335.
5 de Celis JF, Barrio R. Regulation and function of Spalt proteins during animal development. *Int J Dev Biol* 2009; **53**: 1385–1398.
6 Kohlhase J. SALL1 mutations in Townes-Brocks syndrome and related disorders. *Hum Mutat* 2000; **16**: 460–466.
7 Kiefer SM, Robbins L, Stumpff KM, Lin C, Ma L, Rauchman M. Sall1-dependent signals affect Wnt signaling and ureter tip fate to initiate kidney development. *Development* 2010; **137**: 3099–3106.
8 Karantzali E, Lekakis V, Ioannou M, Hadjimichael C, Papamatheakis J, Kretsovali A. Sall1 regulates embryonic stem cell differentiation in association with nanog. *J Biol Chem* 2011; **286**: 1037–1045.
9 Hill VK, Hesson LB, Dansranjavin T, Dallol A, Bieche I, Vacher S *et al.* Identification of 5 novel genes methylated in breast and other epithelial cancers. *Mol Cancer* 2010; **9**: 51.
10 Lee JM, Dedhar S, Kalluri R, Thompson EW. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *J Cell Biol* 2006; **172**: 973–981.
11 Kalluri R. EMT: when epithelial cells decide to become mesenchymal-like cells. *J Clin Invest* 2009; **119**: 1417–1419.
12 May CD, Sphyris N, Evans KW, Werden SJ, Guo W, Mani SA. Epithelial-mesenchymal transition and cancer stem cells: a dangerously dynamic duo in breast cancer progression. *Breast Cancer Res* 2011; **13**: 202.
13 Yang MH, Hsu DS, Wang HW, Wang HJ, Lan HY, Yang WH *et al.* Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition. *Nat Cell Biol* 2010; **12**: 982–992.
14 Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008; **133**: 704–715.
15 Cellecta, http://www.decipherproject.net/ 2012.
16 Netzer C, Rieger L, Brero A, Zhang CD, Hinzke M, Kohlhase J *et al.* SALL1, the gene mutated in Townes-Brocks syndrome, encodes a transcriptional repressor which interacts with TRF1/PIN2 and localizes to pericentromeric heterochromatin. *Hum Mol Genet* 2001; **10**: 3017–3024.
17 Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell* 2009; **139**: 871–890.
18 Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest* 2009; **119**: 1420–1428.
19 Wang CC, Liau JY, Lu YS, Chen JW, Yao YT, Lien HC. Differential expression of moesin in breast cancers and its implication in epithelial-mesenchymal transition. *Histopathology* 2012; **61**: 78–87.
20 Haynes J, Srivastava J, Madson N, Wittmann T, Barber DL. Dynamic actin remodeling during epithelial-mesenchymal transition depends on increased moesin expression. *Mol Biol Cell* 2011; **22**: 4750–4764.
21 Chien W, O'Kelly J, Lu D, Leiter A, Sohn J, Yin D *et al.* Expression of connective tissue growth factor (CTGF/CCN2) in breast cancer cells is associated with increased migration and angiogenesis. *Int J Oncol* 2011; **38**: 1741–1747.
22 Tsai MS, Bogart DF, Castaneda JM, Li P, Lupu R. Cyr61 promotes breast tumorigenesis and cancer progression. *Oncogene* 2002; **21**: 8178–8185.
23 Jing C, El-Ghany MA, Beesley C, Foster CS, Rudland PS, Smith P *et al.* Tazarotene-induced gene 1 (TIG1) expression in prostate carcinomas and its relationship to tumorigenicity. *J Natl Cancer Inst* 2002; **94**: 482–490.
24 Hsu TH, Chu CC, Jiang SY, Hung MW, Ni WC, Lin HE *et al.* Expression of the class II tumor suppressor gene RIG1 is directly regulated by p53 tumor suppressor in cancer cell lines. *FEBS Lett* 2012; **586**: 1287–1293.
25 Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* 2011; **146**: 633–644.
26 Fillmore CM, Kuperwasser C. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res* 2008; **10**: R25.
27 Pawitan Y, Bjohle J, Amler L, Borg AL, Egyhazi S, Hall P *et al.* Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005; **7**: R953–R964.
28 Loi S, Haibe-Kains B, Majjaj S, Lallemand F, Durbecq V, Larsimont D *et al.* PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer. *Proc Natl Acad Sci USA* 2010; **107**: 10208–10213.
29 Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001; **98**: 10869–10874.
30 Ricardo S, Vieira AF, Gerhard R, Leitao D, Pinto R, Cameselle-Teijeiro JF *et al.* Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *J Clin Pathol* 2011; **64**: 937–946.
31 Argos M, Kibriya MG, Jasmine F, Olopade OI, Su T, Hibshoosh H *et al.* Genomewide scan for loss of heterozygosity and chromosomal amplification in breast carcinoma using single-nucleotide polymorphism arrays. *Cancer Genet Cytogenet* 2008; **182**: 69–74.
32 Chin SF, Wang Y, Thorne NP, Teschendorff AE, Pinder SE, Vias M *et al.* Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers. *Oncogene* 2007; **26**: 1959–1970.
33 Onder TT, Gupta PB, Mani SA, Yang J, Lander ES, Weinberg RA. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* 2008; **68**: 3645–3654.
34 Gupta PB, Onder TT, Jiang G, Tao K, Kuperwasser C, Weinberg RA *et al.* Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* 2009; **138**: 645–659.
35 Fredebohm J, Wolf J, Hoheisel JD, Boettcher M. Depletion of RAD17 sensitizes pancreatic cancer cells to gemcitabine. *J Cell Sci* 2013; **126**: 3380–3389.

Supplementary Information accompanies this paper on the Oncogene website (http://www.nature.com/onc)