Published in final edited form as:

Gene. 2016 June 15; 584(2): 111–119. doi:10.1016/j.gene.2016.02.019.

SALL4, the missing link between stem cells, development and cancer

Hiro Tatetsu^{1,§}, Nikki R Kong^{1,§}, Gao Chong¹, Giovanni Amabile², Daniel G Tenen^{3,4}, and Li Chai^{1,‡}

¹Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 77 Avenue Louis Pasteur, New Research Building Room 652D, Boston, MA 02115 ²ADLENNE, Via Zurigo, 46, 6900 Lugano – Switzerland ³Cancer Science Institute of Singapore, National University of Singapore, Centre for Translational Medicine (MD6), #12-01, 14 Medical Drive, Singapore 117599 ⁴Harvard Stem Cell Institute, Center for Life Science Room 437, 3 Blackfan Circle Room 437, Boston, MA 02115

Abstract

There is a growing body of evidence supporting that cancer cells share many similarities with embryonic stem cells (ESCs). For example, aggressive cancers and ESCs share a common gene expression signature that includes hundreds of genes. Since ESC genes are not present in most adult tissues, they could be ideal candidate targets for cancer-specific diagnosis and treatment. This is an exciting cancer-targeting model. The major hurdle to test this model is to identify the key factors/pathway(s) within ESCs that are responsible for the cancer phenotype. SALL4 is one of few genes that can establish this link. The first publication of SALL4 is on its mutation in a human inherited disorder with multiple developmental defects. Since then, over 300 papers have been published on various aspects of this gene in stem cells, development, and cancers. This review aims to summarize our current knowledge of SALL4, including a SALL4-based approach to classify and target cancers. Many questions about this important gene still remain unanswered, specifically, on how this gene regulates cell fates at a molecular level. Understanding SALL4's molecular functions will allow development of specific targeted approaches in the future.

Keywords

SALL4; oncofetal protein

[‡]To whom correspondence should be addressed: L. Chai, lchai@partners.org.

[§]Both authors contribute equally to the work

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest

The authors have no conflict of interest to declare.

The corresponding Gene Wiki entry for this review can be found here: <https://en.wikipedia.org/wiki/SALL4>

1. Introduction

SALL4, a member of the *spalt-like* (SALL) gene family (*SALL1* to *SALL4*), was originally cloned based on its DNA sequence homology to the homeotic gene in *Drosophila*, *spalt* (*sal*)¹²³. *SALL4* plays an essential role in maintaining the pluripotent and self-renewal properties of embryonic stem cells (ESCs). After birth, *SALL4* expression is down-regulated and absent in most adult tissues. However, *SALL4* is re-expressed in various cancers. This review summarizes our current knowledge about the structure and expression patterns of *SALL4*, its role in stem and cancer cells, and the existing understanding of its molecular functions.

1.1. *SALL4* Gene and Protein Structure

In humans, *SALL4* is located in 20q13.2, consisting of four coding exons and 3,162 bp of coding sequence (*SALL4A*; AY172738)⁴⁵. GenBank sequences AY170621 (*SALL4B* mRNA, 1851bp) and AY170622 (*SALL4C* mRNA, 831bp) suggest that two alternative splicing products exist in addition to the full length *SALL4A* mRNA. The B mRNA consists of exon 1, 1,020 bp of 5' end of exon 2, exons 3 and 4, whereas in the C mRNA, exon 2 is spliced out⁶ (Figure 1). Another *SALL4* transcript containing different exons 1a and 1b instead of the original exon 1 identified by 5' RACE has also been reported⁷.

Like the protein encoded by *sal*, *SALL4* belongs to a group of transcription factors characterized by multiple cys(2)his(2) (C2H2)-type zinc finger domains distributed over the entire protein. Human *SALL4* contains a single C2H2 zinc finger near the N-terminus, two C2H2 zinc finger clusters in the middle portion, and one at the C-terminus of the protein⁸ (Figure 1). *SALL4A* has all four zinc finger clusters while *SALL4B* lacks the zinc finger clusters ZF2 and ZF3. The most C-terminal ZF4 cluster of *Sall4* is both necessary and sufficient for its localization to the heterochromatin⁹. In addition to zinc fingers, *SALL4* also contains a glutamine (Q)-rich region that is highly conserved in all invertebrate and vertebrate SALL family members. This domain was shown to be necessary for interactions between Spalt family proteins¹⁰. Furthermore, *SALL4* A and B isoforms are able to form homodimers and heterodimers¹¹, possibly through their Q-rich regions. *SALL4* is a nuclear protein and its subcellular localization is mediated through at least one conserved nuclear localization signal (NLS) at amino acids (AA) 64–67. A single mutation that changes lysine 64 into arginine (K64 into R64) is sufficient to disrupt its subcellular distribution and compromises its function *in vivo*¹².

Post-translational modifications of *SALL4* such as phosphorylation¹³, ubiquitination¹⁴ and sumoylation¹⁵ have been reported. The functions of these post-translational modifications are largely unknown. One study reported that lysine residues 156, 316, 374, and 401 were essential for *SALL4B* sumoylation, which seemed to affect the protein stability, its interaction with Oct4, and its transactivation function¹⁵. The role of ubiquitination of *SALL4* is still unknown.

1.2. SALL4 Expression in Development

In murine development, Sall4 protein expression is first observed at the two-cell stage due to maternal contribution, and later in some cells of the 8- to 16-cell-stage mouse embryo after zygotic transcription has initiated¹⁶. In late blastocysts, the *Sall4* RNA and Sall4 protein become enriched in the inner cell mass (ICM) and the trophectoderm. Finally, by E11.5, Sall4 expression is observed in the midbrain, the rostral edge of the forebrain, maxillary arch, genital tubercle, limb buds, tail, and left ventricular myocardium^{17,18}. In adult mice, Sall4 expression is mostly restricted to germ cells, wherein it is highly expressed in undifferentiated spermatogonia and oocytes in primordial, primary, and secondary follicles^{19,20,21}. Similarly, expression of SALL4 in adult human tissue is restricted to the testis and ovary⁴ (Figure 2). One exception to this expression pattern is human CD34+ hematopoietic stem/progenitor cells (HSPCs)²².

SALL4 has isoform-specific gene expression pattern in the testis, ESCs, and fetal liver cells. *Sall4a* is expressed in postnatal day 7 (PND7) testis, while *Sall4b* is expressed from PND0 onward²³. *Sall4a* is more abundant than *Sall4b* in undifferentiated mouse ESCs, and neither isoform is found upon induction of ESC differentiation¹¹. Furthermore, the A isoform, and not B, is detected in fetal livers, while several human hepatocellular carcinoma (HCC) cell lines express both *SALL4A* and *SALL4B*²⁴.

1.3. SALL4 Gene Regulation

Very little is known about the expression regulation of *SALL4* and its isoforms. Several studies have focused on the promoter region of this gene^{7,25,26}. While one study reported on *SALL4*'s intronic enhancer²⁷, its distal regulatory element(s) remains uninvestigated. In breast cancer, *SALL4* was proposed to be a downstream target gene of STAT3²⁵. During intestinal metaplasia of the gastric epithelial cells, *SALL4* was identified as a direct target of Caudal-related homeobox 1 (CDX1) protein²⁸. In murine ESC, several reports have demonstrated that Sall4 protein participates in an interconnected autoregulatory circuit with Oct4, Sox2, and Nanog, wherein each of the four factors can regulate its own expression as well as that of others^{29,30,31}. Within this circuit, one study reported that Sall4 can negatively self-regulate and antagonize Oct4's activation function to balance its own expression level²⁶. Further, a TCF/LEF consensus sequence was reported in the *SALL4* promoter region and *SALL4* expression could be activated by LEF1 or TCF4E, indicating that *SALL4* is a target of the canonical WNT signaling⁷. Posttranscriptional regulation of *SALL4* has also been observed. Specifically in glioma, one study reported that miR-107 can bind the 3'UTR of *SALL4* mRNA and modulate its expression. In a metastatic breast cancer model³², *SALL4* was found to be a target of miR-33b. In murine ESC, miR-294 and let-7c were reported to have opposing effects on *Sall4* expression³³.

Epigenetic regulation of *SALL4* has also been observed. Specifically, the DNA methylation status of *SALL4* seems to correlate with its expression. In fact, *SALL4* is hypomethylated and highly expressed in induced pluripotent stem (iPS) cells and ESCs, comparing to differentiated cells (Figure 3)^{34,35}. While one study focused on the methylation status of *SALL4* promoter region³⁶, the major CpG island of *SALL4* is located at the Exon 1/Intron 1 region as reported by other *SALL4* methylation studies^{37,38}. Future experiments are thus

needed to explore the effects of upstream regulator(s) on *SALL4* DNA methylation status and how this epigenetic status affects cell fate.

2. Function of *SALL4* in Stem Cells and Development

Sall4 has been shown to be involved in the proper development of the ICM and *Sall4*-null mice do not survive beyond embryonic day E6.5¹⁶⁹³⁰. Although *Sall4*-null blastocysts have no defects in lineage commitment of the ICM or trophoctoderm *in vivo*, *Sall4*-null ICM cells show defected proliferation *in vitro*⁹. Down-regulation of *Sall4* in ESCs via shRNA leads to reduced *Pou5f1* expression and increased *Cdx2* expression, which results in an expansion of trophoctoderm cells³⁰. When cultured with mouse embryonic fibroblast (MEF) feeder cells, one study reported that *Sall4* is essential for efficient proliferation and stabilization, but not for pluripotency, of ESCs by repressing trophoctoderm gene expression³⁹. Another study, however, showed that *Sall4* is essential for the pluripotency of ESCs, especially under feeder-free conditions³⁰. Additional reports suggested that *Sall4* is important for ES cell fate^{26, 27, 30, 31, 40, 41, 42}. *Sall4* was also reported to play a positive role in the generation of iPS cells from fibroblasts⁴³. *SALL4* expression needs to be reactivated during the reprogramming process by which MEF cells are converted to iPS cells⁴⁴. Furthermore, *SALL4* was used to enhance iPS cell reprogramming efficiency in a cell fusion approach⁴⁵, and inclusion of *Sall4* in the Oct4, Sox2, and Klf4 (OSK) reprogramming system leads a more consistent and stable induction of iPS cells with a higher efficiency⁴⁶. Interestingly, ectopic expression of the combination of *Sall4*, *Nanog*, *Esrrb*, and *Lin28* (SNEL), or the combination of *Sall4*, *Sall1*, *Utf1*, *Nanog* and *Myc* in MEFs can reprogram them into iPSCs. These results indicate that *Sall4*, in combination with other factors, can substitute for OSKM factors⁴⁷⁴⁸. Furthermore, *SALL4* can interact with the Nucleosome Remodeling and Histone Deacetylation (NuRD) complex³⁹⁴⁹. MBD3, an essential structural component of the NuRD complex, has been shown to be a key regulator in the reprogramming process⁵⁰⁵¹⁵²⁵³. Though it has been suggested that the *SALL4*/NuRD interaction may not be required in iPSC reprogramming in a cell fusion model⁴⁵, it remains to be seen whether the same conclusion can be drawn from other approaches in generating iPSCs.

Mutations in the *SALL4* gene have been shown to cause autosomal dominant diseases such as Okihiro/Duane-radial ray syndrome (DRRS; MIM 607323)⁴⁵⁴, acro-renal-ocular syndrome (AROS; MIM 102490), Instituto Venezolano de Investigaciones Cientificas syndrome (IVIC)⁵⁵, and are suspected to cause thalidomide embryopathy⁵⁶⁵⁷. The phenotype of Okihiro/Duane-radial ray syndrome has overlap with several other syndromes such as Holt Oram syndrome (HOS; MIM 142900)⁵⁸, which is caused by *TBX5* mutations and Townes-Brocks syndrome, which is caused by *SALL1* mutations (TBS; MIM 107480)⁵⁹.

Okihiro syndrome refers to the familial occurrence of radial-sided hand malformations in association with Duane, a congenital eye movement disorder. The range of associated features among the described families is broad, including anal stenosis, atrial septal defect, hearing impairment, pigmentary disturbance, renal abnormalities, external ear malformations, and facial asymmetry in the affected individuals. Significant intrafamilial variability is a feature of the reported pedigrees⁴. In addition to these developmental defects,

SALL4 was reported as a potential candidate gene that mutated in Han Chinese women with premature ovarian failures⁶⁰.

Homozygous mutations within the *SALL4* gene have not been reported in humans, possibly because this would result in early embryonic lethality. All reported *SALL4* mutations are heterozygous and are located in exons 2 and 3, with no mutations found in exons 1 or 4. These are nonsense mutations, short duplications, or deletions. Two cases of missense *SALL4* mutations around the coding region for the ZF4 cluster were also reported^{61,62}.

In mice, *Sall4* has been reported to be involved in neural tube closure, limb and heart development^{39,63,91,8}. Heterozygous mice carrying a targeted *Sall4* allele with deletion of exons 2–3 that contain all ZF domains exhibit embryonic exencephaly, anorectal malformations, and heart defects⁹. However, in this particular study, limb defects were not described. In a different report, heterozygous mice with a targeted *SALL4* allele that lacks exons 2–4 exhibit deafness, lower anogenital tract abnormalities, anencephaly, Hirschsprung's disease, renal hypoplasia, and skeletal defects⁶³. However, heart anomalies were not described in this model. Mice with a *Sall4* gene trap allele with an insertion in exon 2 exhibit variable embryonic phenotypes ranging from ventricular septum defect (VSD) to mild limb abnormalities¹⁸.

3. *SALL4* in cancers

Analyses of its expression and epigenetic status have shown that *SALL4* is de-regulated and aberrantly expressed in various cancers (Review in⁶⁴) such as leukemia⁶⁵, germ cell tumors^{66,67}, hepatocellular carcinoma (HCC)^{24,68}, gastric cancer^{20,69–73}, colorectal carcinoma^{74–76}, esophageal squamous cell carcinoma^{77,78}, breast cancer^{79–81}, endometrial cancer^{82,83}, lung cancer^{84–86} and glioma⁸⁷. Functionally, *SALL4* is important for the survival^{38,80,82,88–93}, drug resistance^{82,94} and metastasis^{69,82} of many of these cancer cells (Figure 4). The main challenge in this area is the need for standardization of the *SALL4*-based diagnostic methods. While some reports focus on the RNA expression level of *SALL4*, others focus on the protein expression level. Frequently, there is a discrepancy between these two detection methods. Even for *SALL4* protein immunohistochemistry (IHC) studies, various antibodies and different diagnostic criteria were used. Both percentage of positive cells and intensity of staining, as well as the staining pattern, such as nuclear versus cytoplasmic, should be considered. Quality control of *SALL4* antibodies should be evaluated before their use in IHC experiments. Cell type-specific genes, in addition to *SALL4*, should be examined in various *SALL4*-positive cancers. Standardization of *SALL4*-based diagnosis is important for future development of *SALL4*-specific approaches to classify and target cancers. Below, we have detailed selected studies of *SALL4*'s mis-regulation in cancer.

In hematological malignancies, high level of *SALL4* expression is correlated with high-risk myelodysplastic syndromes (MDS) patients with poor survival^{95,96}. Consistent with this observation, aberrant hypomethylation of *SALL4* gene in MDS patients has also been reported^{37,97}. Increased *SALL4* expression was detected in patient samples from blastic stage of chronic myeloid leukemia (CML), compared to the chronic phase, CML patients

who have achieved complete remission (CR) after chemotherapy groups⁹⁰, or those who have tyrosine kinase inhibitor (TKI) resistance⁹⁸. *SALL4* is constitutively expressed in most human primary acute myeloid leukemia (AML) and myeloid leukemia cell lines⁶⁵. Further supporting this, transgenic mice over-expressing *SALL4B* develop hematopoietic disorders, including MDS-like symptoms and AML⁶⁵. *SALL4* is also abnormally expressed in B cell acute lymphocytic leukemia (B-ALL) with hypomethylation at the previously described CpG island^{38,99}. *SALL4* expression is the highest in B-ALL patients with TEL-AML1 translocation, which is the most common genetic abnormality in pediatric B-ALL. In contrast, *SALL4* is negative in T cell acute lymphocytic leukemia (T-ALL), and most lymphomas⁹⁹ with the exception ALK-positive anaplastic large cell lymphoma (ALK+ ALCL)¹⁰⁰.

In solid tumors, *SALL4* was first studied in germ cell tumors, where it is a sensitive diagnostic marker for primary central nerve system (CNS) germ cell tumors⁶⁶, testicular germ cell tumors⁶⁷, primary extragonadal germ cell tumors^{101,102}, mediastinal seminoma^{103,104}, ovarian malignant germ cell tumor¹⁰⁵, and metastatic germ cell tumors¹⁰⁶. In yolk sac tumors and primary extragonadal germ cell tumors, one study reported that the circulating tumor cells are detectable by anti-*SALL4* antibody in peripheral blood samples, making this protein a more sensitive marker than α -fetoprotein and glypican-3¹⁰⁷. In addition, *SALL4* is expressed in a majority of malignant rhabdoid tumors (MRTs)²⁰¹⁰⁸¹⁰⁹¹¹⁰¹¹¹ but rarely expressed in epithelioid sarcomas. Therefore, *SALL4* immunoexpression may be a useful diagnostic tool to distinguish epithelioid sarcoma from malignant rhabdoid tumor.

For tumors in the gastrointestinal & digestive system, *SALL4* is well studied in HCC. *SALL4* is expressed in human fetal livers and absent in adult livers²⁴⁶⁸, but it is re-expressed in a subgroup of HCC patients. *SALL4*-positive HCC cells have a gene expression pattern that is similar to that of fetal hepatic progenitor cells. These HCCs tend to be more aggressive and associated with a poor prognosis⁶⁸. Compared to *SALL4*-negative cases, HCC patients with *SALL4* expression show higher frequency of tumor cell invasion, intrahepatic metastasis, and are positive for cytokeratin 19 and Epithelial Cell Adhesion Molecule (EpCAM), as well as other cancer stem cell markers¹¹². A positive correlation between *SALL4* and EpCAM expression have been reported in other HCC studies⁹¹¹¹³¹¹⁴¹¹⁵. Liver cancer patients with high serum level of *SALL4* have poor prognosis as evidenced by both tumor recurrence and overall survival rate¹¹⁶.

4. Molecular dissection of *SALL4* functions in cells

To understand the mechanism(s) underlying *SALL4* function(s) in various cells, efforts have been made to search for both *SALL4* protein interaction partners and its downstream targets. The PTEN/AKT pathway is the best-understood example that integrates our current understanding of these two aspects of *SALL4*'s molecular functions. Based on chromatin immunoprecipitation followed by microarray (ChIP-chip) data in normal human CD34+ cells, leukemic NB4 cells, and human ESCs, *PTEN* was identified to be a key *SALL4* downstream target gene^{49,92}. *SALL4* directly interacts with the NuRD complex, and recruits it to the promoter region of *PTEN*⁴⁹. A peptide that competes with *SALL4* in its interaction

with the NuRD complex can reverse its effect on *PTEN* repression, which leads to the activation of the AKT pathway^{88,93}. Here we have summarized some of the known interacting partners of SALL4 and its downstream target genes.

4.1. Protein-partners of SALL4

It has been shown that SALL4 can interact with many proteins, which can define its molecular function. Specifically, SALL4 can bind the NuRD complex, LSD1, or DNMTs, thus functions as a transcriptional repressor. SALL4 can also form a protein complex with transcriptional regulators such as Mixed Lineage Leukemia (MLL) or β -catenin to activate genes. Furthermore, Sall4 interacts with Oct4, Nanog, or Sox2 to maintain “stemness” of ESCs. There are further reports of SALL4 interaction with TBX5 in cardiac tissues¹⁸, with PLZF in spermatogonial progenitors in postnatal testes¹¹⁷, and with Rad50 to stabilize the Mre11–Rad50–Nbs1 complex during the DNA damage repair process¹¹⁸. We have detailed some of these interactions below.

A conserved N-terminal 12 amino acid (N-12aa) domain of SALL4 that is also present in other SALL family members can recruit the NuRD complex and mediate transcriptional repression^{119,120}. The NuRD complex comprises many different subunits, the most frequently associated ones are histone deacetylases HDAC1/2, ATP-dependent remodeling enzymes CHD3/4, histone chaperones RbAp46/48, CpG-binding proteins MBD2/3, the GATAD2a (p66a) and/or GATAD2b (p66b), and sequence-specific DNA-binding proteins MTA1/2/3^{121–123}. SALL4’s recruitment of the NuRD complex was validated by studies of a SALL4 target gene, *PTEN*, which encodes an important tumor repressor that regulates the AKT signaling pathway and cell growth. SALL4 co-occupies the *PTEN* promoter region with a NuRD component, HDAC2, to repress its gene expression⁴⁹. Treatment of SALL4-expressing malignant cells with a synthesized N-12aa peptide to compete for SALL4’s interaction with NuRD leads to cell death due to *PTEN* de-repression, which can be rescued by co-treatment of cells with a PTEN inhibitor^{88,93}.

SALL4’s interaction with NuRD may mediate its binding to the NuRD-associated LSD1 protein, a histone demethylase that specifically converts histone H3K4me2 to H3K4me1 or unmethylated H3K4, thus functioning as a transcriptional co-repressor^{124,125}. Both WT SALL4A and SALL4B proteins can co-immunoprecipitate (IP) with LSD1¹²⁶. Based on serial truncations of SALL4 in co-IP and demethylation experiments, the region of SALL4 containing the Q-rich domain is essential for its binding to LSD1. Both gain- and loss-of-function approaches revealed that SALL4 controls the histone binding levels of LSD1, which is accompanied by changes in the methylation status of H3K4 at promoter regions of genes that they co-regulate¹²⁶. In addition to its function in modulating the covalent modification of histones, there is also evidence that SALL4 may directly interact with DNA methyltransferases (DNMTs) to mediate gene repression³⁶. Furthermore, the C-terminal portion of SALL4 can bind Cyclin D1, which together with SALL4, can associate with heterochromatic region and modulate gene repression¹²⁷.

Through interacting with other proteins, SALL4 can also activate gene expression. For example, SALL4 recruits the MLL complex, which contains histone methyltransferase activity for K4 and K79 on the histone 3 tail. In the context of AML, this interaction results

in *HOXA9* up-regulation and subsequently more aggressive leukemogenesis⁸⁹. Members of the MLL complex, including RbBp5 and Menin, were also found in SALL4 pull-down experiments⁸⁹. In addition, MLL complex and SALL4 are known to mediate signaling from the Wnt/ β -catenin pathway in AML development. Wnt is a highly conserved pathway that plays an important role in development and cancer¹²⁸. In amphibian neural development, Sall4 is known to be activated by Wnt signaling to specify posterior neural fate¹²⁹. Furthermore, both SALL4A and SALL4B isoforms can bind β -catenin protein and synergistically activate the Wnt/ β -catenin signaling pathway⁶⁵. The interaction domain between SALL4 and β -catenin was not identified. However, based on binding data between SALL1 and β -catenin, the C-terminal portion of SALL4 is likely involved in this interaction¹³⁰. Elucidating the relationship among SALL4, MLL, and Wnt signaling can lead to novel targeting opportunities in AML treatment.

Many of the interactions described above between SALL4 and its co-regulators were observed in the context of leukemogenesis or cancer development. Due to its key roles in ESCs, SALL4's interactions with transcription factors crucial for pluripotency have also been explored. Two independent studies, using unbiased approaches of affinity purification of Oct4 followed by mass spectrometry, identified Sall4 to be one of its binding partners in mouse ESCs^{131,132}. Both approaches identified components of the NuRD complex, raising the possibility that Sall4 may bridge the interaction between Oct4 and this epigenetic repressor complex. Since Oct4, Sox2, and Nanog form a central pluripotency transcription factor network in ESCs, Sall4's interaction with these proteins was also investigated. Through affinity precipitation using oligonucleotides containing OCT-SOX elements, Sall4 was found to be specifically associated with Oct4/Sox2/DNA complex in mouse ESCs. Though it is unclear which part of Sall4 is responsible for binding to Sox2 or whether the interaction is direct, this association is found at the promoter of a subset of pluripotent genes⁴². Another pluripotency factor, Nanog, also seem to be present in the same protein complex with Sall4 in ESCs, with the N-terminal portion of Sall4 important for this interaction²⁷. As is the case in regards to Sox2, it is unclear whether the interaction between Nanog and Sall4 is direct. Moreover, a recent study suggested that SALL4 can recruit polycomb complex members such as Bmi-1 to the *Sox* gene locus, thus facilitating changes in higher-order chromatin structure¹³³. These studies further support that Sall4 is part of the important transcription regulatory complex that is responsible for maintaining the pluripotency of mouse ESCs.

SALL4 may play a role in the DNA damage response. Sall4 is required for efficient activation of ATM-dependent responses to DNA double stranded breaks (DSBs) damage in ESCs. Sall4 was observed to rapidly mobilize to the sites of DSBs after DNA damage. Identified by mass-spec and confirmed by co-IP, Sall4 interacts with Rad50 and stabilizes the Mre11–Rad50–Nbs1 complex for the efficient recruitment and activation of ATM. Further, Sall4 interacts with Baf60a, a member of the SWI/SNF (switch/sucrose nonfermentable) ATP-dependent chromatin-remodeling complex, which may be responsible for recruiting Sall4 to the site of DNA DSB¹¹⁸.

4.2 Downstream targets and pathways

The DNA binding ability of SALL4 remains to be resolved. Genome-wide analyses such as ChIP followed by next-generation sequencing (seq) or ChIP-chip have suggested several potential binding motifs for SALL4. The common SALL4 DNA binding motif suggested by several groups^{11,134,135} is “ATTTGCAT”, which appears similar to the consensus motif utilized by multiple factors such as Nanog, Oct4, and Sox2^{135–139}. A study on differential DNA binding potential of SALL4 isoforms¹¹ suggested that another motif, “TTGTCTACTTGTA”, exists for Sall4A, and was predominantly found in differentiation and patterning genes. A different study reported Sall4 ChIP-chip data that revealed an enrichment of the sequence “TCGCCATA” in Sall4 peaks with the highest enrichment³¹. Further studies are needed to verify which motif is important using known SALL4 downstream target genes.

Many critical downstream targets of SALL4 have been identified. SALL4 can bind its own promoter region and exerts a strong self-repressive auto-regulation, which in turn acts as a “gate keeper” for the SALL4/OCT4 positive feedback loop²⁶. SALL4 can also regulate other ESC genes such as *Nanog*, *Pou5f1*, and *SOX2*. In addition, in normal human hematopoietic and leukemic cells, SALL4 can up-regulate the expression of *BMI-1*^{95,98,140}, *HOXA*^{92,89}, and down-regulate *PTEN* expression⁴⁹. Being a survival factor for cancer cells, SALL4 can affect apoptosis-inducing genes (e.g., *TNF*, *TP53*, *PTEN*, *CARD9*, *CARD11*, *CYCS*, *LTA*) and apoptosis-inhibiting genes (e.g., *BMI-1*, *BCL2*, *XIAP*, *DAD1*, *TEGT*)⁹². *MYC* has been shown to be a SALL4 target both in cancer (endometrial cancer) and ESCs^{29,82,92}. SALL4 is enriched in side population (SP) cells and can regulate the expression of ATP-binding cassette multidrug transporter, *ABCA3* and *ABCB1*⁹⁴. Genome-wide analysis revealed that SALL4 regulates the Sonic Hedgehog (SHH) pathway²⁹. Sall4 and Gli3 cooperate for proper development of the anterior-proximal skeletal elements and also function upstream of Sonic hedgehog (Shh)-dependent posterior skeletal element development¹⁴¹. Finally, SALL4 can regulate the expression of cell surface molecules such as EpCAM^{91,113}, epithelial-mesenchymal transition (EMT) related factors (e.g. *SNAI1*¹¹³, *CXCR4*, *TWIST1*⁹¹, E-Cadherin^{69,83,142}, *VIM*, and *ZEB1*¹⁴³) in cancers (Figure 5).

5. Conclusion remarks

The fate of a specific cell type is governed by genetic and epigenetic factors, and can be redirected or reprogrammed through modifications of the key factors involved. Identifying and understanding the key players in cell fate determination will help us direct a cell type-specific gene expression profile, which is ultimately responsible for cellular identity and function. The knowledge thus gained can help us direct cellular differentiation for the purpose of tissue regeneration and repair, or control malignant cell growth. The zinc-finger protein SALL4 has a unique role in connecting ESCs, development, and cancer. SALL4 affects its downstream target gene expression by interacting with various epigenetic modulators and sequence-specific transcription factors, thus making it a key cell fate regulator. Additional studies are needed to shed light on the molecular mechanism(s) of SALL4 functions in various cell types in order to open avenues for new therapeutic

approaches to target cancers as well as generate stem-cell related applications for tissue regeneration.

Acknowledgments

This review and the corresponding Gene Wiki article are written as part of the Gene Wiki Review series—a series resulting from a collaboration between the journal GENE and the Gene Wiki Initiative. The Gene Wiki Initiative is supported by National Institutes of Health (GM089820). Additional support for Gene Wiki Reviews is provided by Elsevier, the publisher of GENE. This work is also supported in part through NIH grant PO1HL095489, research funds from Leukemia and Lymphoma Society and V foundation (to LC), and by the Singapore Ministry of Health's National Medical Research Council under its Singapore Translational Research (STaR) Investigator Award and the National Research Foundation Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiative (to DGT). All authors contribute to the writings of the review. We want to thank Alicia J Stein for reviewing the manuscript.

References

1. Kohlhase J, et al. Isolation, characterization, and organ-specific expression of two novel human zinc finger genes related to the Drosophila gene spalt. *Genomics*. 1996; 38:291–298. [PubMed: 8975705]
2. Frei E, et al. Molecular characterization of spalt, a homeotic gene required for head and tail development in the Drosophila embryo. *The EMBO journal*. 1988; 7:197–204. [PubMed: 16453821]
3. Kuhnlein RP, et al. spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. *The EMBO journal*. 1994; 13:168–179. [PubMed: 7905822]
4. Kohlhase J, et al. Okihiro syndrome is caused by SALL4 mutations. *Human molecular genetics*. 2002; 11:2979–2987. [PubMed: 12393809]
5. Deloukas P, et al. The DNA sequence and comparative analysis of human chromosome 20. *Nature*. 2001; 414:865–871. [PubMed: 11780052]
6. Kohlhase J, et al. SALL4 mutations in Okihiro syndrome (Duane-radial ray syndrome), acrorenalocular syndrome, and related disorders. *Human mutation*. 2005; 26:176–183. [PubMed: 16086360]
7. Bohm J, Sustmann C, Wilhelm C, Kohlhase J. SALL4 is directly activated by TCF/LEF in the canonical Wnt signaling pathway. *Biochem Biophys Res Commun*. 2006; 348:898–907. [PubMed: 16899215]
8. de Celis JF, Barrio R. Regulation and function of Spalt proteins during animal development. *Int J Dev Biol*. 2009; 53:1385–1398. [PubMed: 19247946]
9. Sakaki-Yumoto M, et al. The murine homolog of SALL4, a causative gene in Okihiro syndrome, is essential for embryonic stem cell proliferation, and cooperates with Sall1 in anorectal, heart, brain and kidney development. *Development*. 2006; 133:3005–3013. [PubMed: 16790473]
10. Sweetman D, Smith T, Farrell ER, Chantry A, Munsterberg A. The conserved glutamine-rich region of chick csal1 and csal3 mediates protein interactions with other spalt family members. Implications for Townes-Brocks syndrome. *The Journal of biological chemistry*. 2003; 278:6560–6566. [PubMed: 12482848]
11. Rao S, et al. Differential roles of Sall4 isoforms in embryonic stem cell pluripotency. *Mol Cell Biol*. 2010; 30:5364–5380. [PubMed: 20837710]
12. Wu M, et al. Identification of the nuclear localization signal of SALL4B, a stem cell transcription factor. *Cell cycle*. 2014; 13:1456–1462. [PubMed: 24626181]
13. Zoumaro-Djayoon AD, et al. Investigating the role of FGF-2 in stem cell maintenance by global phosphoproteomics profiling. *Proteomics*. 2011; 11:3962–3971. [PubMed: 21761559]
14. Wilson CH, et al. Nuclear receptor binding protein 1 regulates intestinal progenitor cell homeostasis and tumour formation. *The EMBO journal*. 2012; 31:2486–2497. [PubMed: 22510880]

15. Yang F, et al. Sumoylation is important for stability, subcellular localization, and transcriptional activity of SALL4, an essential stem cell transcription factor. *The Journal of biological chemistry*. 2012; 287:38600–38608. [PubMed: 23012367]
16. Elling U, Klasen C, Eisenberger T, Anlag K, Treier M. Murine inner cell mass-derived lineages depend on Sall4 function. *Proc Natl Acad Sci U S A*. 2006; 103:16319–16324. [PubMed: 17060609]
17. Kohlhasse J, et al. Cloning and expression analysis of SALL4, the murine homologue of the gene mutated in Okihiro syndrome. *Cytogenet Genome Res*. 2002; 98:274–277. [PubMed: 12826753]
18. Koshiba-Takeuchi K, et al. Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart. *Nature genetics*. 2006; 38:175–183. [PubMed: 16380715]
19. Eildermann K, et al. Developmental expression of the pluripotency factor sal-like protein 4 in the monkey, human and mouse testis: restriction to premeiotic germ cells. *Cells, tissues, organs*. 2012; 196:206–220. [PubMed: 22572102]
20. Miettinen M, et al. SALL4 expression in germ cell and non-germ cell tumors: a systematic immunohistochemical study of 3215 cases. *The American journal of surgical pathology*. 2014; 38:410–420. [PubMed: 24525512]
21. Cao D, Guo S, Allan RW, Molberg KH, Peng Y. SALL4 is a novel sensitive and specific marker of ovarian primitive germ cell tumors and is particularly useful in distinguishing yolk sac tumor from clear cell carcinoma. *The American journal of surgical pathology*. 2009; 33:894–904. [PubMed: 19295406]
22. Gao C, et al. SALL4 is a key transcription regulator in normal human hematopoiesis. *Transfusion*. 2013; 53:1037–1049. [PubMed: 22934838]
23. Gassei K, Orwig KE. SALL4 expression in gonocytes and spermatogonial clones of postnatal mouse testes. *PLoS One*. 2013; 8:e53976. [PubMed: 23326552]
24. Oikawa T, et al. Sall4 regulates cell fate decision in fetal hepatic stem/progenitor cells. *Gastroenterology*. 2009; 136:1000–1011. [PubMed: 19185577]
25. Bard JD, et al. Signal transducer and activator of transcription 3 is a transcriptional factor regulating the gene expression of SALL4. *FASEB J*. 2009; 23:1405–1414. [PubMed: 19151334]
26. Yang J, Gao C, Chai L, Ma Y. A novel SALL4/OCT4 transcriptional feedback network for pluripotency of embryonic stem cells. *PLoS One*. 2010; 5:e10766. [PubMed: 20505821]
27. Wu Q, et al. Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. *The Journal of biological chemistry*. 2006; 281:24090–24094. [PubMed: 16840789]
28. Fujii Y, et al. CDX1 confers intestinal phenotype on gastric epithelial cells via induction of stemness-associated reprogramming factors SALL4 and KLF5. *Proc Natl Acad Sci U S A*. 2012; 109:20584–20589. [PubMed: 23112162]
29. Yang J, et al. Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine embryonic stem cells. *Proc Natl Acad Sci U S A*. 2008; 105:19756–19761. [PubMed: 19060217]
30. Zhang J, et al. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol*. 2006; 8:1114–1123. [PubMed: 16980957]
31. Lim CY, et al. Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. *Cell Stem Cell*. 2008; 3:543–554. [PubMed: 18804426]
32. Lin Y, et al. MicroRNA-33b Inhibits Breast Cancer Metastasis by Targeting HMGA2, SALL4 and Twist1. *Sci Rep*. 2015; 5:9995. [PubMed: 25919570]
33. Melton C, Judson RL, Blueloch R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature*. 2010; 463:621–626. [PubMed: 20054295]
34. Nishino K, et al. Defining hypo-methylated regions of stem cell-specific promoters in human iPS cells derived from extra-embryonic amnions and lung fibroblasts. *PLoS One*. 2010; 5:e13017. [PubMed: 20885964]
35. Amabile G, et al. Dissecting the role of aberrant DNA methylation in human leukaemia. *Nat Commun*. 2015; 6:7091. [PubMed: 25997600]

36. Yang J, Corsello TR, Ma Y. Stem cell gene SALL4 suppresses transcription through recruitment of DNA methyltransferases. *The Journal of biological chemistry*. 2012; 287:1996–2005. [PubMed: 22128185]
37. Lin J, et al. Aberrant hypomethylation of SALL4 gene in patients with myelodysplastic syndrome. *Leukemia research*. 2013; 37:71–75. [PubMed: 23122807]
38. Ueno S, et al. Aberrant expression of SALL4 in acute B cell lymphoblastic leukemia: mechanism, function, and implication for a potential novel therapeutic target. *Experimental hematology*. 2014; 42:307–316. e308. [PubMed: 24463278]
39. Yuri S, et al. Sall4 is essential for stabilization, but not for pluripotency, of embryonic stem cells by repressing aberrant trophectoderm gene expression. *Stem cells*. 2009; 27:796–805. [PubMed: 19350679]
40. Wang J, et al. A protein interaction network for pluripotency of embryonic stem cells. *Nature*. 2006; 444:364–368. [PubMed: 17093407]
41. Zhou Q, Chipperfield H, Melton DA, Wong WH. A gene regulatory network in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*. 2007; 104:16438–16443. [PubMed: 17940043]
42. Tanimura N, Saito M, Ebisuya M, Nishida E, Ishikawa F. Stemness-related factor Sall4 interacts with transcription factors Oct-3/4 and Sox2 and occupies Oct-Sox elements in mouse embryonic stem cells. *The Journal of biological chemistry*. 2013; 288:5027–5038. [PubMed: 23269686]
43. Tsubooka N, et al. Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. *Genes Cells*. 2009; 14:683–694. [PubMed: 19476507]
44. Hou P, et al. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science*. 2013; 341:651–654. [PubMed: 23868920]
45. Wong CC, Gaspar-Maia A, Ramalho-Santos M, Reijo Pera RA. High-efficiency stem cell fusion-mediated assay reveals Sall4 as an enhancer of reprogramming. *PLoS One*. 2008; 3:e1955. [PubMed: 18414659]
46. Parchem RJ, et al. Two miRNA clusters reveal alternative paths in late-stage reprogramming. *Cell Stem Cell*. 2014; 14:617–631. [PubMed: 24630794]
47. Buganim Y, et al. The developmental potential of iPSCs is greatly influenced by reprogramming factor selection. *Cell Stem Cell*. 2014; 15:295–309. [PubMed: 25192464]
48. Mansour AA, et al. The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature*. 2012; 488:409–413. [PubMed: 22801502]
49. Lu J, et al. Stem cell factor SALL4 represses the transcriptions of PTEN and SALL1 through an epigenetic repressor complex. *PLoS One*. 2009; 4:e5577. [PubMed: 19440552]
50. Rais Y, et al. Deterministic direct reprogramming of somatic cells to pluripotency. *Nature*. 2013; 502:65–70. [PubMed: 24048479]
51. Lujan E, et al. Early reprogramming regulators identified by prospective isolation and mass cytometry. *Nature*. 2015; 521:352–356. [PubMed: 25830878]
52. Luo M, et al. NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. *Stem Cells*. 2013; 31:1278–1286. [PubMed: 23533168]
53. Ozmadenci D, et al. Netrin-1 regulates somatic cell reprogramming and pluripotency maintenance. *Nat Commun*. 2015; 6:7398. [PubMed: 26154507]
54. Al-Baradie R, et al. Duane radial ray syndrome (Okihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. *American journal of human genetics*. 2002; 71:1195–1199. [PubMed: 12395297]
55. Paradisi I, Arias S. IVIC syndrome is caused by a c.2607delA mutation in the SALL4 locus. *Am J Med Genet A*. 2007; 143:326–332. [PubMed: 17256792]
56. Kohlhasse J, et al. Mutations at the SALL4 locus on chromosome 20 result in a range of clinically overlapping phenotypes, including Okihiro syndrome, Holt-Oram syndrome, acro-renal-ocular syndrome, and patients previously reported to represent thalidomide embryopathy. *Journal of medical genetics*. 2003; 40:473–478. [PubMed: 12843316]
57. Kohlhasse J, Holmes LB. Mutations in SALL4 in malformed father and daughter postulated previously due to reflect mutagenesis by thalidomide. *Birth defects research. Part A, Clinical and molecular teratology*. 2004; 70:550–551.

58. Basson CT, et al. The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). *The New England journal of medicine*. 1994; 330:885–891. [PubMed: 8114858]
59. Kohlhase J, Wischermann A, Reichenbach H, Froster U, Engel W. Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. *Nat Genet*. 1998; 18:81–83. [PubMed: 9425907]
60. Wang B, et al. Mutational analysis of SAL-Like 4 (SALL4) in Han Chinese women with premature ovarian failure. *Molecular human reproduction*. 2009; 15:557–562. [PubMed: 19581335]
61. Miertus J, et al. A SALL4 zinc finger missense mutation predicted to result in increased DNA binding affinity is associated with cranial midline defects and mild features of Okihiro syndrome. *Human genetics*. 2006; 119:154–161. [PubMed: 16402211]
62. Diehl A, Mu W, Batista D, Gunay-Aygun M. An atypical 0.73 MB microduplication of 22q11.21 and a novel SALL4 missense mutation associated with thumb agenesis and radioulnar synostosis. *Am J Med Genet A*. 2015; 167:1644–1649. [PubMed: 25823593]
63. Warren M, et al. A Sall4 mutant mouse model useful for studying the role of Sall4 in early embryonic development and organogenesis. *Genesis*. 2007; 45:51–58. [PubMed: 17216607]
64. Zhang X, Yuan X, Zhu W, Qian H, Xu W. SALL4: an emerging cancer biomarker and target. *Cancer Lett*. 2015; 357:55–62. [PubMed: 25444934]
65. Ma Y, et al. SALL4, a novel oncogene, is constitutively expressed in human acute myeloid leukemia (AML) and induces AML in transgenic mice. *Blood*. 2006; 108:2726–2735. [PubMed: 16763212]
66. Mei K, et al. Diagnostic utility of SALL4 in primary germ cell tumors of the central nervous system: a study of 77 cases. *Mod Pathol*. 2009; 22:1628–1636. [PubMed: 19820689]
67. Cao D, et al. RNA-binding protein LIN28 is a marker for testicular germ cell tumors. *Hum Pathol*. 2011; 42:710–718. [PubMed: 21288558]
68. Yong KJ, Chai L, Tenen DG. Oncofetal gene SALL4 in aggressive hepatocellular carcinoma. *The New England journal of medicine*. 2013; 369:1171–1172. [PubMed: 24047068]
69. Zhang L, et al. SALL4, a novel marker for human gastric carcinogenesis and metastasis. *Oncogene*. 2014; 33:5491–5500. [PubMed: 24276240]
70. Ushiku T, et al. SALL4 represents fetal gut differentiation of gastric cancer, and is diagnostically useful in distinguishing hepatoid gastric carcinoma from hepatocellular carcinoma. *The American journal of surgical pathology*. 2010; 34:533–540. [PubMed: 20182341]
71. Ikeda H, et al. alpha-Fetoprotein-producing gastric carcinoma and combined hepatocellular and cholangiocarcinoma show similar morphology but different histogenesis with respect to SALL4 expression. *Hum Pathol*. 2012; 43:1955–1963. [PubMed: 22516245]
72. Osada M, et al. Combination of hepatocellular markers is useful for prognostication in gastric hepatoid adenocarcinoma. *Hum Pathol*. 2014; 45:1243–1250. [PubMed: 24767858]
73. Sugai T, et al. Molecular analysis of gastric differentiated-type intramucosal and submucosal cancers. *International journal of cancer. Journal international du cancer*. 2010; 127:2500–2509. [PubMed: 20178104]
74. Forghanifard MM, et al. Role of SALL4 in the progression and metastasis of colorectal cancer. *J Biomed Sci*. 2013; 20:6. [PubMed: 23363002]
75. Ardalan Kholes S, et al. SALL4 as a new biomarker for early colorectal cancers. *J Cancer Res Clin Oncol*. 2015; 141:229–235. [PubMed: 25156818]
76. Habano W, Sugai T, Jiao YF, Nakamura S. Novel approach for detecting global epigenetic alterations associated with tumor cell aneuploidy. *International journal of cancer. Journal international du cancer*. 2007; 121:1487–1493. [PubMed: 17546590]
77. Singhi AD, et al. Undifferentiated carcinoma of the esophagus: a clinicopathological study of 16 cases. *Hum Pathol*. 2015; 46:366–375. [PubMed: 25582499]
78. Forghanifard MM, et al. Stemness state regulators SALL4 and SOX2 are involved in progression and invasiveness of esophageal squamous cell carcinoma. *Med Oncol*. 2014; 31:922. [PubMed: 24659265]
79. Yue X, et al. High cytoplasmic expression of SALL4 predicts a malignant phenotype and poor prognosis of breast invasive ductal carcinoma. *Neoplasma*. 2015

80. Kobayashi D, Kuribayashi K, Tanaka M, Watanabe N. SALL4 is essential for cancer cell proliferation and is overexpressed at early clinical stages in breast cancer. *Int J Oncol.* 2011; 38:933–939. [PubMed: 21274508]
81. Panis C, Pizzatti L, Herrera AC, Cecchini R, Abdelhay E. Putative circulating markers of the early and advanced stages of breast cancer identified by high-resolution label-free proteomics. *Cancer Lett.* 2013; 330:57–66. [PubMed: 23200674]
82. Li A, et al. SALL4 is a new target in endometrial cancer. *Oncogene.* 2015; 34:63–72. [PubMed: 24336327]
83. Liu L, et al. SALL4 as an Epithelial-Mesenchymal Transition and Drug Resistance Inducer through the Regulation of c-Myc in Endometrial Cancer. *PLoS One.* 2015; 10:e0138515. [PubMed: 26407074]
84. Kobayashi D, Kuribayashi K, Tanaka M, Watanabe N. Overexpression of SALL4 in lung cancer and its importance in cell proliferation. *Oncol Rep.* 2011; 26:965–970. [PubMed: 21725617]
85. Fujimoto M, et al. SALL4 immunohistochemistry in non-small-cell lung carcinomas. *Histopathology.* 2014; 64:309–311. [PubMed: 24117842]
86. Morita S, et al. High-grade lung adenocarcinoma with fetal lung-like morphology: clinicopathologic, immunohistochemical, and molecular analyses of 17 cases. *The American journal of surgical pathology.* 2013; 37:924–932. [PubMed: 23629442]
87. Zhang L, et al. The expression of SALL4 in patients with gliomas: high level of SALL4 expression is correlated with poor outcome. *J Neurooncol.* 2015; 121:261–268. [PubMed: 25359397]
88. Gao C, et al. Targeting transcription factor SALL4 in acute myeloid leukemia by interrupting its interaction with an epigenetic complex. *Blood.* 2013; 121:1413–1421. [PubMed: 23287862]
89. Li A, et al. A SALL4/MLL/HOXA9 pathway in murine and human myeloid leukemogenesis. *The Journal of clinical investigation.* 2013; 123:4195–4207. [PubMed: 24051379]
90. Lu J, et al. Dissecting the role of SALL4, a newly identified stem cell factor, in chronic myelogenous leukemia. *Leukemia.* 2011; 25:1211–1213. [PubMed: 21468036]
91. Oikawa T, et al. Sal-like protein 4 (SALL4), a stem cell biomarker in liver cancers. *Hepatology.* 2013; 57:1469–1483. [PubMed: 23175232]
92. Yang J, et al. SALL4 is a key regulator of survival and apoptosis in human leukemic cells. *Blood.* 2008; 112:805–813. [PubMed: 18487508]
93. Yong KJ, et al. Oncofetal gene SALL4 in aggressive hepatocellular carcinoma. *The New England journal of medicine.* 2013; 368:2266–2276. [PubMed: 23758232]
94. Jeong HW, et al. SALL4, a stem cell factor, affects the side population by regulation of the ATP-binding cassette drug transport genes. *PLoS One.* 2011; 6:e18372. [PubMed: 21526180]
95. Wang F, et al. Stem cell factor SALL4, a potential prognostic marker for myelodysplastic syndromes. *J Hematol Oncol.* 2013; 6:73. [PubMed: 24283704]
96. Shuai X, et al. Overexpression of the novel oncogene SALL4 and activation of the Wnt/beta-catenin pathway in myelodysplastic syndromes. *Cancer Genet Cytogenet.* 2009; 194:119–124. [PubMed: 19781444]
97. Ma JC, et al. Aberrant hypomethylation of SALL4 gene is associated with intermediate and poor karyotypes in acute myeloid leukemia. *Clinical biochemistry.* 2013; 46:304–307. [PubMed: 23201546]
98. Shen Q, et al. The differential expression pattern of the BMI-1, SALL4 and ABCA3 genes in myeloid leukemia. *Cancer Cell Int.* 2012; 12:42. [PubMed: 23067006]
99. Cui W, et al. Differential expression of the novel oncogene, SALL4, in lymphoma, plasma cell myeloma, and acute lymphoblastic leukemia. *Mod Pathol.* 2006; 19:1585–1592. [PubMed: 16998462]
100. Wang P, et al. The expression and oncogenic effects of the embryonic stem cell marker SALL4 in ALK-positive anaplastic large cell lymphoma. *Cellular signalling.* 2012; 24:1955–1963. [PubMed: 22743134]
101. Cao D, et al. RNA-binding protein LIN28 is a marker for primary extragonadal germ cell tumors: an immunohistochemical study of 131 cases. *Mod Pathol.* 2011; 24:288–296. [PubMed: 21057460]

102. Camparo P, Comperat EM. SALL4 is a useful marker in the diagnostic work-up of germ cell tumors in extra-testicular locations. *Virchows Arch.* 2013; 462:337–341. [PubMed: 23224120]
103. Iwamoto N, et al. Mediastinal seminoma: a case report with special emphasis on SALL4 as a new immunocytochemical marker. *Diagn Cytopathol.* 2013; 41:821–824. [PubMed: 22298374]
104. Weissferdt A, et al. Primary mediastinal seminomas: a comprehensive immunohistochemical study with a focus on novel markers. *Hum Pathol.* 2015; 46:376–383. [PubMed: 25576290]
105. He L, et al. Overexpression of karyopherin 2 in human ovarian malignant germ cell tumor correlates with poor prognosis. *PLoS One.* 2012; 7:e42992. [PubMed: 22962582]
106. Cao D, Humphrey PA, Allan RW. SALL4 is a novel sensitive and specific marker for metastatic germ cell tumors, with particular utility in detection of metastatic yolk sac tumors. *Cancer.* 2009; 115:2640–2651. [PubMed: 19365862]
107. Nastaly P, et al. Circulating tumor cells in patients with testicular germ cell tumors. *Clin Cancer Res.* 2014; 20:3830–3841. [PubMed: 24634372]
108. Deisch J, Raisanen J, Rakheja D. Immunohistochemical expression of embryonic stem cell markers in malignant rhabdoid tumors. *Pediatr Dev Pathol.* 2011; 14:353–359. [PubMed: 21417895]
109. Venneti S, et al. Malignant rhabdoid tumors express stem cell factors, which relate to the expression of EZH2 and Id proteins. *The American journal of surgical pathology.* 2011; 35:1463–1472. [PubMed: 21921784]
110. Yoshida A, et al. Differential SALL4 immunoexpression in malignant rhabdoid tumours and epithelioid sarcomas. *Histopathology.* 2015; 66:252–261. [PubMed: 24827994]
111. Kohashi K, et al. ERG and SALL4 expressions in SMARCB1/INI1-deficient tumors: a useful tool for distinguishing epithelioid sarcoma from malignant rhabdoid tumor. *Hum Pathol.* 2015; 46:225–230. [PubMed: 25479928]
112. Shibahara J, et al. Clinicopathologic characteristics of SALL4-immunopositive hepatocellular carcinoma. *Springerplus.* 2014; 3:721. [PubMed: 26034695]
113. Zeng SS, et al. The transcription factor SALL4 regulates stemness of EpCAM-positive hepatocellular carcinoma. *J Hepatol.* 2014; 60:127–134. [PubMed: 24012616]
114. Park H, et al. SALL4 Expression in Hepatocellular Carcinomas Is Associated with EpCAM-Positivity and a Poor Prognosis. *J Pathol Transl Med.* 2015; 49:373–381. [PubMed: 26265684]
115. Tanaka Y, et al. Spalt-like transcription factor 4 immunopositivity is associated with epithelial cell adhesion molecule expression in combined hepatocellular carcinoma and cholangiocarcinoma. *Histopathology.* 2015
116. Han SX, et al. Serum SALL4 is a novel prognosis biomarker with tumor recurrence and poor survival of patients in hepatocellular carcinoma. *J Immunol Res.* 2014; 2014:262385. [PubMed: 24860834]
117. Hobbs RM, et al. Functional antagonism between Sall4 and Plzf defines germline progenitors. *Cell Stem Cell.* 2012; 10:284–298. [PubMed: 22385656]
118. Xiong J, et al. Stemness factor Sall4 is required for DNA damage response in embryonic stem cells. *J Cell Biol.* 2015; 208:513–520. [PubMed: 25733712]
119. Hong W, et al. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *The EMBO journal.* 2005; 24:2367–2378. [PubMed: 15920470]
120. Lauberth SM, Rauchman M. A conserved 12-amino acid motif in Sall1 recruits the nucleosome remodeling and deacetylase corepressor complex. *The Journal of biological chemistry.* 2006; 281:23922–23931. [PubMed: 16707490]
121. Lai AY, Wade PA. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nature reviews. Cancer.* 2011; 11:588–596. [PubMed: 21734722]
122. Torchy MP, Hamiche A, Klaholz BP. Structure and function insights into the NuRD chromatin remodeling complex. *Cellular and molecular life sciences : CMLS.* 2015; 72:2491–2507. [PubMed: 25796366]
123. Allen HF, Wade PA, Kutateladze TG. The NuRD architecture. *Cellular and molecular life sciences : CMLS.* 2013; 70:3513–3524. [PubMed: 23340908]

124. Shi Y, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*. 2004; 119:941–953. [PubMed: 15620353]
125. Forneris F, Binda C, Vanoni MA, Mattevi A, Battaglioli E. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS letters*. 2005; 579:2203–2207. [PubMed: 15811342]
126. Liu L, et al. Histone lysine-specific demethylase 1 (LSD1) protein is involved in Sal-like protein 4 (SALL4)-mediated transcriptional repression in hematopoietic stem cells. *The Journal of biological chemistry*. 2013; 288:34719–34728. [PubMed: 24163373]
127. Bohm J, Kaiser FJ, Borozdin W, Depping R, Kohlhasse J. Synergistic cooperation of Sall4 and Cyclin D1 in transcriptional repression. *Biochem Biophys Res Commun*. 2007; 356:773–779. [PubMed: 17383611]
128. Wang Y, et al. The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science*. 2010; 327:1650–1653. [PubMed: 20339075]
129. Young JJ, Kjolby RA, Kong NR, Monica SD, Harland RM. Spalt-like 4 promotes posterior neural fates via repression of pou5f3 family members in *Xenopus*. *Development*. 2014; 141:1683–1693. [PubMed: 24715458]
130. Sato A, et al. Sall1, a causative gene for Townes-Brocks syndrome, enhances the canonical Wnt signaling by localizing to heterochromatin. *Biochem Biophys Res Commun*. 2004; 319:103–113. [PubMed: 15158448]
131. van den Berg DL, et al. An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell*. 2010; 6:369–381. [PubMed: 20362541]
132. Li L, et al. Synergistic transcriptional and post-transcriptional regulation of ESC characteristics by core pluripotency transcription factors in protein-protein interaction networks. *PLoS One*. 2014; 9:e105180. [PubMed: 25171496]
133. Abboud N, et al. A cohesin-OCT4 complex mediates Sox enhancers to prime an early embryonic lineage. *Nat Commun*. 2015; 6:6749. [PubMed: 25851587]
134. Paik EJ, et al. A Cdx4-Sall4 regulatory module controls the transition from mesoderm formation to embryonic hematopoiesis. *Stem Cell Reports*. 2013; 1:425–436. [PubMed: 24286030]
135. Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell*. 2008; 132:1049–1061. [PubMed: 18358816]
136. Bailey TL, Williams N, Misleh C, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic acids research*. 2006; 34:W369–W373. [PubMed: 16845028]
137. Loh YH, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature genetics*. 2006; 38:431–440. [PubMed: 16518401]
138. Macisaac KD, et al. A hypothesis-based approach for identifying the binding specificity of regulatory proteins from chromatin immunoprecipitation data. *Bioinformatics*. 2006; 22:423–429. [PubMed: 16332710]
139. Tsankov AM, et al. Transcription factor binding dynamics during human ES cell differentiation. *Nature*. 2015; 518:344–349. [PubMed: 25693565]
140. Yang J, et al. Bmi-1 is a target gene for SALL4 in hematopoietic and leukemic cells. *Proc Natl Acad Sci U S A*. 2007; 104:10494–10499. [PubMed: 17557835]
141. Akiyama R, et al. Sall4-Gli3 system in early limb progenitors is essential for the development of limb skeletal elements. *Proc Natl Acad Sci U S A*. 2015; 112:5075–5080. [PubMed: 25848055]
142. Deng G, et al. SALL4 is a novel therapeutic target in intrahepatic cholangiocarcinoma. *Oncotarget*. 2015; 6:27416–27426. [PubMed: 26317546]
143. Ito J, Matsumoto Y, Yoshikawa K, Toi M. Sal-like 4 (SALL4) suppresses CDH1 expression and maintains cell dispersion in basal-like breast cancer. *FEBS letters*. 2013; 587:3115–3121. [PubMed: 23954296]

Highlights

This review summarizes the gene structure, expression and function(s) of SALL4.

SALL4 plays an essential role in embryonic stem cells.

Mutations in the *SALL4* gene causes Okihiro/Duane-radial ray syndrome.

SALL4 is de-regulated and aberrantly expressed in various cancers.

SALL4 can be a cancer-specific target.

Figure 1. Human *SALL4* gene structure and isoforms

Alternative splicing generates 3 forms of *SALL4* mRNA. *SALL4A* (encoding 1053 amino acids) has all 4 exons and contains 4 zinc finger domains (ZF1–4, indicated by blue bars). *SALL4B* (encoding 617 a.a.) is generated via an alternative splice donor site that results in the deletion of large portion of exon 2 and a protein that contains only ZF1 and ZF4. *SALL4C* (encoding 277 a.a.) does not have exon 2 and encodes a version that only has ZF1. A conserved N-terminal 12 amino acid (N-12aa) domain of *SALL4* (indicated by red bars) is sufficient for recruiting the nucleosome remodeling and histone deacetylase (NuRD) complex and mediating transcriptional repression. A glutamine (Q)-rich region (indicated by yellow bars) that is highly conserved in all invertebrate and vertebrate *Spalt* family members, was shown to be necessary in protein interactions between *Spalt* family proteins. A conserved nuclear localization signal (NLS) has been reported and is indicated by green bars.

Figure 2. SALL4 expression during development

Sall4 protein expression is first observed at the two-cell stage, and later in some cells of the 8- to 16-cell-stages of the mouse embryo. In late blastocysts, Sall4 protein becomes enriched in the inner cell mass (ICM) and the trophectoderm. By E11.5, Sall4 expression is observed in the midbrain, genital tubercle, limb and tail buds^{17,18}. In adult mice, Sall4 expression is mostly restricted to the germ cells, wherein it is highly expressed in undifferentiated spermatogonia and oocytes in primordial, primary, and secondary follicles^{19,20,21}. The red circles represent SALL4 expression.

Figure 3. SALL4 CpG island
SALL4 DNA methylation changes at the single CpG level during reprogramming of K562 cells to leukemia-derived pluripotent stem cells (LiPS) cells. Results from K562 and two separate LiPS lines are shown³⁵.

Figure 4. SALL4 expression in cancer

SALL4 is de-regulated and aberrantly expressed in various cancers.

Figure 5. Summary of regulation and downstream targets of SALL4

Many critical interacting partners, upstream regulators, as well as downstream targets of SALL4 are summarized here. The red arrow indicates activation and the blue bar indicates repression.