Overexpression of Sal-like protein 4 in head and neck cancer: epigenetic effects and clinical correlations

Overexpression of Sal-like protein 4 in head and neck cancer: Epigenetic effects and clinical correlations

Kiyoshi Misawa* [a], Yuki Misawa [a], Masato Mima [a, b], Satoshi Yamada [a, c], Atsushi Imai [a], Daiki Mochizuki [a], Takuya Nakagawa [b], Tomoya Kurokawa [b], Shiori Endo [a], Hideya Kawasaki [c], Brenner J. Chad [d] and Hiroyuki Mineta [a]

[a] Department of Otolaryngology/Head and Neck Surgery, Hamamatsu University School of Medicine, Shizuoka, Japan

[b] Department of Molecular Oncology, Chiba University, Japan

[c] Preeminent Medical Photonics Education & Research Center Institute for NanoSuit Research, Hamamatsu University School of Medicine, Hamamatsu, Japan

[d] Department of Otolaryngology-Head and Neck Surgery, University of Michigan Health System, Ann Arbor, MI, USA.

*** Corresponding author:** Kiyoshi Misawa, Department of Otolaryngology/Head and Neck Surgery, 1-20-1 Handayama, Hamamatsu University School of Medicine, Shizuoka 431-3192, Japan; Phone: 81-53-435-2252, Fax: 81-53-435-2253; E-mail: kiyoshim@hama-med.ac.jp

Running title: Overexpression of Sal-like protein 4 in head and neck cancer

Abbreviations: SALL4: Sal-like protein 4; HNSCC: head and neck squamous cell carcinoma; DNMT3A: DNA methyltransferase 3 alpha; DFS: disease-free survival; ROC:

Receiver operating characteristic; TCGA: The Cancer Genome Atlas; EGFR: epidermal growth factor receptor; qRT-PCR: quantitative reverse transcription PCR; Q-MSP: quantitative methylation-specific PCR

5

Abstract

Background Sal-like protein 4 (SALL4), an embryonic stem cell factor, has been reported to play an essential role in embryogenesis and oncogenesis. As yet, however, the expression and role of this transcription factor in head and neck squamous cell carcinoma (HNSCC) has not been established.

Methods We assessed *SALL4* mRNA expression in a well-characterised dataset of 230 HNSCC samples (test cohort 110 cases and validation cohort 120 cases). We also transfected HNSCC cells (FaDu and UM-SCC-6) with SALL4 siRNA and assessed its effects on proliferation and expression of specific epigenetic factors in order to uncover the role of *SALL4* in HNSCC.

Results Overexpression of *SALL4* was detected in tumour samples of both cohorts. HNSCC cells treated with SALL4 siRNA showed a reduction in growth and a decrease in DNA methyltransferase 3 alpha (*DNMT3A*) expression. In the patient cohorts, *SALL4* overexpression was found to significantly correlate with disease recurrence ($p < 0.001$) and *SALL4* methylation status ($p = 0.002$). We also found that *DNMT3A* was significantly upregulated upon *SALL4* upregulation ($p < 0.001$). High expression levels of *SALL4* correlated with decreases in disease-free survival (DFS) rates (log-rank test, $p < 0.001$). Multivariate analysis revealed that *SALL4* expression served as an independent prognostic factor for DFS (hazard ratio: 2.566, 95% confidence interval: 1.598–4.121; $p < 0.001$).

Conclusions Our findings indicate that *SALL4* upregulation correlates with HNSCC tumour aggressiveness and an adverse patient outcome. Our findings also indicate that DNMT3A may synergistically contribute to the regulatory effects of SALL4. Our findings provide insight into SALL4-mediated HNSCC development via epigenetic modulation.

Keywords: HNSCC, SALL4, Biomarker, Epigenetic regulation, DNMT3A

1 Introduction

Patients with head and neck squamous cell carcinoma (HNSCC) require multimodal therapy including surgery, radiation therapy, chemotherapy and/or targeted therapy [1-4]. Despite recent advances in these therapeutic regimens, only slight improvements in survival have been made [5, 6]. HNSCC is a heterogeneous multifactorial disease involving both genetic and epigenetic events [7-9]. Therefore, a refined molecular characterisation is required to identify novel prognostic markers and therapeutic targets.

Sal-like protein 4 (SALL4) is a zinc finger transcription factor that has been reported to be essential for embryonic development and pluripotency [10]. In addition, *SALL4* has been found to represent a genetic bridge between stem cells and malignant cells [11]. Although its levels are downregulated or absent in most adult tissues, *SALL4* has been found to be highly expressed in various human tumours [12] and, as such, to serve as a diagnostic marker with adequate sensitivity [13]. Concordantly, high *SALL4* expression has e.g. been found to be associated with a worse prognosis in digestive tract cancers [14]. Recent studies revealed that SALL4 may also act as epigenetic regulator [15] and that its overexpression may induce hypermethylation of promoter regions of target genes [16].

Our recent efforts to determine the methylation profiles of *SALL1*, *SALL2* and *SALL3* confirmed that their hypermethylation is common in cancer and is associated with disease-free survival (DFS) [17-19]. *SALL1* hypermethylation was significantly correlated with a reduced DFS in patients with early stage T1 and T2 HNSCC [18] and *SALL2* hypermethylation was significantly correlated with a reduced DFS in patients with oral cancer [19]. Furthermore, *SALL3* was found to be associated with aberrant methylation of other tumour-related genes and, as such, to induce critical events in HNSCC progression [17].

The subsequent analysis of *SALL4* expression levels and methylation patterns of tumour-related genes is considered to be important for both understanding HNSCC

development and the design of novel targeted therapies. Here, we aimed to investigate the role of *SALL4* regulation in HNSCC. In addition, we aimed to assess the expression status of *SALL4*-associated genes in HNSCCs and to evaluate their clinical significance.

2 Materials and methods

2.1 Tumour samples

Tissues samples (n = 230) were obtained from patients that underwent surgical resection for HNSCC. All patients provided written informed consent and the study protocol was approved by the Institutional Review Board of the Hamamatsu University School of Medicine (Hamamatsu, Shizuoka, Japan). Two independent cohorts (test and validation) consisted of 110 and 120 fresh frozen samples, respectively. Detailed clinical information was obtained from the patients' medical records (Supplementary Table S1).

2.2 Cell culture

The FaDu cell line, derived from a hypopharyngeal tumour, was purchased from the American Type Culture Collection (ATTC; Manassas, VA, USA). The UM-SCC-6 cell line, derived from a primary tumour of the base of the tongue, was obtained from the University of Michigan (Ann Arbor, MI, USA). The cells were cultured in Dulbecco's modified Eagle's medium (DMEM; FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) supplemented with 10% foetal bovine serum (FBS; Gibco, ThermoFisher Scientific Inc., Waltham, MA, USA) and 1% penicillin/streptomycin (Wako) in a humidified atmosphere containing 5% CO_2 at 37°C.

2.3 siRNA delivery and colony forming assays

SALL4 siRNA (si-SALL4) (SALLL4HSS183741; Invitrogen Thermo Fisher Scientific, CA, USA) was diluted in Opti-MEM I Reduced Serum Medium (Invitrogen) and Lipofectamin RNAiMAX Transfection Reagent (Invitrogen) according to the manufacturer's instructions. Stealth RNAi Negative Control (si-NC) in Medium GC Duplex #2 (Invitrogen) was used as a negative control. The siRNA complexes were delivered to cells at a final concentration of 50 nM and incubated for 2-3 days for RNA and protein assays, and delivered at 1 nM and incubated for 14 days for colony forming assays. To this end, approximately 500 cells were seeded into each well of 6-well plates and allowed to grow for 14 days, after which the cells were fixed and stained with 40% ethanol-1% crystal violet. Cell colonies with diameters > 0.1 mm were enumerated.

2.4 RNA extraction and qRT-PCR

Total RNA was isolated from tissues using a RNeasy Plus Mini Kit (Qiagen, Hilden, Germany) after which complementary DNA (cDNA) was synthesised using a ReverTra Ace qPCR RT Kit (Toyobo, Tokyo, Japan). qRT-PCR assays were performed as described previously [20]. The primer sets used are listed in Supplementary Table S2. The results were analysed using the $^{\Delta\Delta}$Ct method.

2.5 Western blotting

Cells were lysed using RIPA Buffer (Wako) containing a Protease Inhibitor Cocktail (Promega, Madison, WI, USA). The supernatants were collected and the protein concentrations were measured using a Protein Assay Rapid Kit WAKO II (Wako) and standardised against β-Actin. Next, the proteins were resolved using SDS-PAGE with 7.5% TGX FastCast Acrylamide Kit gels (Bio-Rad Laboratories Inc., Hercules, CA, USA) and transferred to Immobilon-P PVDF membranes (Merck KgaA, Darmstadt, Germany). The resulting blots were incubated overnight

with an anti-SALL4 antibody (clone: EE-30, sc-101147; Santa Cruz Biotechnology Inc., TX, USA) and an anti-β-Actin antibody (cat. no. A2228; Sigma-Aldrich, Merck) at 4°C, followed by incubation with a biotin-conjugated secondary anti-mouse antibody (Nichirei Biosciences Inc., Tokyo, Japan) and a horseradish peroxidase-conjugated avidin-biotin reaction (Nichirei). Immunoreactive bands were visualised using an enhanced chemiluminescence substrate (GE Healthcare UK Ltd, Buckinghamshire, England). Western blot images were captured using a ChemiDoc Touch imaging system (Bio-Rad) and analysed using Image Lab 5.2 (Bio-Rad) software.

2.6 Bisulphite treatment and quantitative methylation-specific PCR (Q-MSP) analysis

DNA extraction and bisulphite modification of genomic DNA were carried out as described previously [18]. DNA methylation at CpG sites near promoter regions of the target genes was assessed by Q-MSP using a Thermal Cycler Dice Real Time System TP800 (TaKaRa Bio Inc., Shiga, Japan) [21]. The primer sequences used are listed in the Supplementary Table S2. The methylation density values in the individual samples were determined by calculating gene methylation rates [21].

2.7 Collection of publicly available datasets

Gene expression data were obtained from the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) and MethHC (http://methhc.mbc.nctu.edu.tw/php/index.php) in July 2019 [22]. Expression data were obtained as processed RNA-seq data in the form of RNA-seq via expectation maximisation.

2.8 Data analysis and statistics

SALL4 expression status results and patient characteristics were compared using Student's *t*-test. Receiver operating characteristic (ROC) curve analysis was performed using normalised methylation values of 120 HNSCC and 120 adjacent normal mucosal samples in conjunction with a Stata/SE 13.0 system (Stata Corporation, TX, USA). The area under the ROC curve indicated optimal sensitivity and specificity cut-off values for distinguishing between gene expression levels in HNSCC and normal tissues (Fig. 1C).

DFS was measured from the date of the initial treatment to the date of diagnosis of locoregional recurrence or distant metastasis. Survival probabilities were calculated using the Kaplan-Meier method, and survival rates were compared using the log-rank test. The prognostic value of *SALL4* expression status was assessed using multivariate Cox proportional hazards analysis adjusted for age (≥ 65 versus < 65 years), sex, alcohol intake, smoking status and tumour stage (I, II and III versus IV). All statistical analyses were performed using the StatMate IV software tool (ATMS Co. Ltd., Tokyo, Japan). Differences with $p < 0.05$ were considered statistically significant.

3 Results

3.1 *SALL4* expression levels are increased in HNSCC versus normal mucosal tissues

SALL4 mRNA expression levels were determined in 110 HNSCC (test cohort) and 73 adjacent normal mucosal tissues. The *SALL4* mRNA levels were found to be significantly higher in the HNSCC tissues than in the adjacent normal tissues ($p = 0.002$) (Fig. 1A). Furthermore, upon comparing *SALL4* expression in another cohort of 120 pairs of HNSCC (validation cohort) and normal tissues, the *SALL4* mRNA levels were found to be 3-fold higher in the HNSCC tissues than in the paired non-cancerous mucosal tissues ($p < 0.001$) (Fig. 1B). The *SALL4* mRNA

expression level data generated highly discriminative ROC curve profiles, which clearly distinguished HNSCC from normal mucosal tissues (area under the ROC = 0.771). At the cut-off value of 7.438, the sensitivity and specificity were 79.87% and 65.77%, respectively (Fig. 1C).

3.2 SALL4 knockdown inhibits colony formation and downregulates DNMT3A expression

Next, we performed in vitro experiments to investigate the effect of siRNA-mediated SALL4 knockdown using SALL4-expressing HNSCC cell lines FaDu and UM-SCC-6. We found that SALL4 siRNA significantly downregulated SALL4 protein and mRNA levels, as indicated by Western blot ($p < 0.05$) (Fig. 2A) and qRT-PCR analyses ($p < 0.05$) (Fig. 2B), respectively. In addition, we found that FaDu and UM-SCC-6 cells in which SALL4 was downregulated showed significantly reduced colony forming abilities ($p < 0.05$) (Fig. 2C). Compared to the control group (si-NC), we also found that the mRNA expression level of the DNA methyltransferase *DNMT3A* in the si-SALL4 group was significantly downregulated $p < 0.05$) (Fig. 2D), while no significant difference in the expression of *DNMT3B* was observed (Fig. 2E).

3.3 Correlation between *SALL4* expression and methylation levels in HNSCC tissues

Eleven tumour suppressor genes were defined as methylated genes in each sample (Fig. 3A, Supplementary Table S3). The mean differences in the gene methylation rates of the 11 tumour suppressor genes as determined based on *SALL4* gene expression are illustrated in Fig. 3B. In particular, we found that the gene methylation rates were significantly higher in patients showing *SALL4* upregulation (49.7 ± 19.9%) than in those showing *SALL4* downregulation (41.0 ± 19.0%) ($p = 0.002$) (Fig. 3B). Next, we assessed the *DNMT3A* and *DNMT3B* mRNA levels in the HNSCC specimens using qRT-PCR. Subsequent Spearman's correlation analysis

revealed a positive correlation between *SALL4* and *DNMT3A* (R^2 = 0.4848, $p < 0.001$) expression, while no correlation was observed between *SALL4* and *DNMT3B* (R^2 = 0.0045, p = 0.397) expression (Fig. 3C, 3D). Also, no significant association between *SALL4* expression and *SALL1*, *SALL2* or *SALL3* methylation status was observed (Supplementary Table S4).

3.4 Correlation between *SALL4* expression and clinicopathological characteristics

The major clinicopathological characteristics of the patients are listed in Table 1. In the test cohort, the *SALL4* mRNA levels were found to be significantly associated with disease recurrence (p = 0.006). The *SALL4* expression levels in patients in the validation cohort were significantly associated with smoking status (p = 0.004) and recurrence (p = 0.020). When the test and validation cohorts were merged, *SALL4* expression was significantly correlated with recurrence ($p < 0.001$) (Table 1). Site-specific *SALL4* expression levels in the oral cavity (n = 76), hypopharynx (n = 54), larynx (n = 51), oropharynx (n = 35) and paranasal cavity (n = 14) are listed in Supplementary Table S5.

3.5 Prognostic value of *SALL4* expression

In the test cohort, a shorter DFS was also observed when *SALL4* upregulation was noted compared to *SALL4* downregulation (log-rank test, p = 0.001) (Fig. 4A). The *SALL4* gene was also upregulated in patients in the validation cohort exhibiting a shorter DFS compared to those showing *SALL4* downregulation (log-rank test, p = 0.012) (Fig. 4B). When the test and validation sets were combined, the DFS rate among patients showing *SALL4* upregulation was 38.9% compared to 69.5% among those showing *SALL4* downregulation (log-rank test, $p < 0.001$) (Fig. 4C). Moreover, among 105 patients with tumour stage T1 or T2, the DFS rate among those showing *SALL4* gene upregulation was 40.1%, compared to 72.3% among those showing *SALL4* downregulation (log-rank test, p = 0.009) (Fig. 4D). Among 106 patients

without lymph node metastasis, those with *SALL4* upregulation exhibited a shorter DFS than those with *SALL4* downregulation (log-rank test, $p = 0.002$) (Fig. 4E). Among the 61 patients with stage I or II HNSCC, the DFS rate was lower in the *SALL4* upregulation group compared to the *SALL4* downregulation group (log-rank test, $p = 0.003$) (Fig. 4F). No significant difference in site-specific DFS time among patients with *SALL4* upregulation compared to those with *SALL4* downregulation was observed, with one notable exception, i.e., it was significantly shorter when *SALL4* was upregulated in oral cavity cancer (log-rank test, $p = 0.026$) (Supplementary Fig. S1).

In addition, the association between *SALL4* expression and the risk of recurrence was estimated via multivariate analysis using a Cox proportional hazards model adjusted for age, sex, alcohol exposure, smoking status and clinical stage. Among patients showing *SALL4* upregulation (135/230, 58.7%), the adjusted odds ratio for recurrence was 2.566 (95% confidence interval: 1.598–4.121, $p < 0.001$) (Table 2).

3.6 Independent validation using TCGA data

TCGA data from 43 normal tissue samples and 497 HNSCC tissue samples were analysed. The HNSCC samples showed significantly higher *SALL4* expression levels than the non-tumorous tissues ($p < 0.001$) (Supplementary Fig. S2A). Spearman's correlation analysis revealed that *SALL4* expression was positively correlated with *DNMT3A* expression ($R^2 = 0.0581$, $p < 0.001$) (Supplementary Fig. S2B). No correlation between *SALL4* expression and *DNMT3A* expression was noted ($R^2 = 0.00003$, $p = 0.909$) (Supplementary Fig. S2C). We also found that the *SALL4* mRNA level was inversely correlated with that of *CDH1* (R2 = 0.008, $p = 0.044$) (Supplementary Fig. S3A), and positively with that of *COL1A2* and *CDH13* (R2 = 0.332, $p < 0.001$ and R2 = 0.018, $p = 0.002$, respectively) (Supplementary Fig. S3B, S3C).

4 Discussion

Our data indicate that *SALL4* mRNA expression can be detected with high sensitivity and specificity in HNSCC tissues and that its upregulation is associated with an increased recurrence risk. In addition, we found that siRNA-based downregulation of *SALL4* in HNSCC cells resulted in growth inhibition, supporting the hypothesis that *SALL4* may act as an oncogene. Epigenetic regulation of the *SALL4* gene may play a regulatory role in HNSCC tumorigenesis and in disease recurrence. From our results we conclude that *SALL4* may serve as a putative therapeutic target in HNSCC.

SALL4 has been reported to sustain the stemness of embryonic stem cells by establishing a regulatory signal transduction network with Sox2, Nanog and Oct4 [12]. SALL4 has also been found to be required for the proliferation, invasion and malignant transformation of stem cells present in leukaemia, breast cancer, colon cancer and liver cancer [11, 23-25]. Upregulation of SALL4 expression through epidermal growth factor receptor (EGFR) activation can decrease the sensitivity of EGFR tyrosine kinase inhibitors in CD44-positive lung cancer [26]. Upregulated SALL4 expression levels, which are indicative of stemness-driven tumours with activation of the PTEN-PI3K-AKT pathway and a high histone deacetylase activity, have been found to predict aggressiveness in e.g. hepatocellular carcinoma [27-29]. *SALL4* knockdown has been found to increase the expression of the cell-cell adhesion gene *CDH1* and to reduce the levels of mesenchymal genes in basal-like breast cancer [30]. Conversely, *SALL4* expression has been reported to induce an epithelial-mesenchymal transition (EMT) phenotype and to play an important role in the maintenance of cancer stem cell-like properties such as drug resistance [31].

SALL4 targets specific DNA methyltransferases and, thus, may bring about specific DNA methylation patterns of target genes in various cell systems [15]. SALL4 has also been shown to interact with the histone H3 lysine 36 trimethylation-specific methyltransferase nuclear receptor binding SET domain protein 2, which affects histone modification and, by doing so, regulates the expression of its target genes [32-34]. In addition, SALL4 has been found to repress the expression of PTEN and SALL1 by interacting with the nucleosome remodelling and deacetylase complex [35]. The dynamic regulation of SALL4-associated epigenetic factors may provide promising opportunities for the management of head and neck cancer.

Previous studies have shown that *SALL4* is aberrantly expressed in a variety of human cancer types and is significantly associated with a poor prognosis [29, 36, 37]. Initially, SALL4 was found to be involved in leukemogenesis, exhibiting overexpression in human acute myeloid leukaemia [38]. Subsequently, *SALL4* overexpression was found to serve as an independent prognostic marker and even as a therapeutic target for patients with breast cancer, glioma, endometrial carcinoma, lung cancer, germ cell cancer and hepatoblastoma [39-44]. Although *SALL4* has gained interest as a marker of aggressive subgroups in several cancers, its clinical importance in HNSCC has remained unknown. Our current results implicate *SALL4* expression and DNA promoter methylation of tumour suppressor genes in the genesis of HNSCC. SALL4 may also serve as a novel therapeutic target for the treatment of HNSCC. The expression of SALL4 has been found to be associated with *EGFR* mutations and, consequently, SALL4 inhibition was found to increase the sensitivity to EGFR inhibitors (EGFR tyrosine kinase inhibitors) in CD44-positive lung cancers [26]. SALL4 also promotes gastric cancer metastasis through activation of the TGF-β/SMAD signalling pathway and induction of EMT, indicating that SALL4 may serve as a target for the treatment of gastric cancer [45]. In hepatocellular carcinoma SALL4, which is re-expressed through hepatitis B virus-induced

STAT3 activation, counteracts miR-200c in PD-L1-induced T cell exhaustion which, in turn, improves its clinical outcome [23]. To improve the survival rate of HNSCC patients, novel less toxic treatment strategies are required. Our current results may be instrumental for the development of new drugs and the identification of novel biomarkers.

Acknowledgments: The authors would like to thank Ms. Yuko Mohri for her excellent technical support and Editage (www.editage.jp) for English language editing.

Conflict of Interest: None to declare

Grant support: This study was funded by a Grant-in-Aid for Scientific Research (No. 17K11380, No. 17K16903, 17K16904, 19K09866, 19K09906 and 19K18728) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Author contributions

KM and YM conceived the study. KM, YM and KH designed the experiments. MM, YS, SI, DM, TK, TK, SE and MM analysed the data and prepared the figures and tables. All authors participated in writing the manuscript, reviewed its drafts, approved its final version and agreed with its submission.

Ethics approval and consent to participate

The research methodology employed in this study was approved by The Institutional Review Board of the Hamamatsu University School of Medicine. All study subjects provided written informed consent.

Consent for publication

Consent for publication was obtained from all patients.

5

References

1 M.R. Migden, D. Rischin, C.D. Schmults, A. Guminski, A. Hauschild, K.D. Lewis, C.H. Chung, L. Hernandez-Aya, A.M. Lim, A.L.S. Chang, G. Rabinowits, A.A. Thai, L.A. Dunn, B.G.M. Hughes, N.I. Khushalani, B. Modi, D. Schadendorf, B. Gao, F. Seebach, S. Li, J. Li, M. Mathias, J. Booth, K. Mohan, E. Stankevich, H.M. Babiker, I. Brana, M. Gil-Martin, J. Homsi, M.L. Johnson, V. Moreno, J. Niu, T.K. Owonikoko, K.P. Papadopoulos, G.D. Yancopoulos, I. Lowy and M.G. Fury, PD-1 Blockade with Cemiplimab in Advanced Cutaneous Squamous-Cell Carcinoma, N Engl J Med 379, 341-351 (2018)

2 T. Kanazawa, K. Misawa, Y. Misawa, T. Uehara, H. Fukushima, G. Kusaka, M. Maruta and T.E. Carey, G-Protein-Coupled Receptors: Next Generation Therapeutic Targets in Head and Neck Cancer?, Toxins (Basel) 7, 2959-2984 (2015)

3 M. Mikulandra, A. Kobescak, B. Verillaud, P. Busson and T. Matijevic Glavan, Radio-sensitization of head and neck cancer cells by a combination of poly(I:C) and cisplatin through downregulation of survivin and c-IAP2, Cell Oncol (Dordr) 42, 29-40 (2019)

4 S. Roy, M. Kar, A. Saha, S. Padhi and B. Banerjee, Role of beta-catenin in cisplatin resistance, relapse and prognosis of head and neck squamous cell carcinoma, Cell Oncol (Dordr) 41, 185-200 (2018)

5 V. Budach and I. Tinhofer, Novel prognostic clinical factors and biomarkers for outcome prediction in head and neck cancer: a systematic review, Lancet Oncol 20, e313-e326 (2019)

6 J.Y. Hur, H.R. Kim, J.Y. Lee, S. Park, J.A. Hwang, W.S. Kim, S. Yoon, C.M. Choi, J.K. Rho and J.C. Lee, CDK7 inhibition as a promising therapeutic strategy for lung squamous cell carcinomas with a SOX2 amplification, Cell Oncol (Dordr) 42, 449-458 (2019)

7 P. Boscolo-Rizzo, M.C. Da Mosto, E. Rampazzo, S. Giunco, A. Del Mistro, A. Menegaldo, L. Baboci, M. Mantovani, G. Tirelli and A. De Rossi, Telomeres and telomerase in head and neck squamous cell carcinoma: from pathogenesis to clinical implications, Cancer Metastasis Rev 35, 457-474 (2016)

8 K. Misawa, Y. Misawa, T. Kanazawa, D. Mochizuki, A. Imai, S. Endo, T.E. Carey and H. Mineta, Epigenetic inactivation of galanin and GALR1/2 is associated with early recurrence in head and neck cancer, Clin Exp Metastasis 33, 187-195 (2016)

9 T. Nakagawa, K. Matsusaka, K. Misawa, S. Ota, K. Takane, M. Fukuyo, B. Rahmutulla, K.I. Shinohara, N. Kunii, D. Sakurai, T. Hanazawa, H. Matsubara, Y. Nakatani, Y. Okamoto and A. Kaneda, Frequent promoter hypermethylation associated with human papillomavirus infection in pharyngeal cancer, Cancer Lett 407, 21-31 (2017)

10 J. Kohlhase, R. Schuh, G. Dowe, R.P. Kuhnlein, H. Jackle, B. Schroeder, W. Schulz-Schaeffer, H.A. Kretzschmar, A. Kohler, U. Muller, M. Raab-Vetter, E. Burkhardt, W. Engel and R. Stick, Isolation, characterization, and organ-specific expression of two novel human zinc finger genes related to the Drosophila gene spalt, Genomics 38, 291-298 (1996)

11 L. Yang, L. Liu, H. Gao, J.P. Pinnamaneni, D. Sanagasetti, V.P. Singh, K. Wang, M. Mathison, Q. Zhang, F. Chen, Q. Mo, T. Rosengart and J. Yang, The stem cell factor SALL4 is an essential transcriptional regulator in mixed lineage leukemia-rearranged leukemogenesis, J Hematol Oncol 10, 159 (2017)

12 H. Tatetsu, N.R. Kong, G. Chong, G. Amabile, D.G. Tenen and L. Chai, SALL4, the missing link between stem cells, development and cancer, Gene 584, 111-119 (2016)

13 K. Mei, A. Liu, R.W. Allan, P. Wang, Z. Lane, T.W. Abel, L. Wei, H. Cheng, S. Guo, Y. Peng, D. Rakheja, M. Wang, J. Ma, M.M. Rodriguez, J. Li and D. Cao, Diagnostic utility of SALL4 in primary germ cell tumors of the central nervous system: a study of 77 cases, Mod Pathol 22, 1628-1636 (2009)

14 K.J. Yong, C. Gao, J.S. Lim, B. Yan, H. Yang, T. Dimitrov, A. Kawasaki, C.W. Ong, K.F. Wong, S. Lee, S. Ravikumar, S. Srivastava, X. Tian, R.T. Poon, S.T. Fan, J.M. Luk, Y.Y. Dan, M. Salto-Tellez, L. Chai and D.G. Tenen, Oncofetal gene SALL4 in aggressive hepatocellular carcinoma, N Engl J Med 368, 2266-2276 (2013)

15 J. Yang, T.R. Corsello and Y. Ma, Stem cell gene SALL4 suppresses transcription through recruitment of DNA methyltransferases, J Biol Chem 287, 1996-2005 (2012)

16 J. Xiong, Z. Zhang, J. Chen, H. Huang, Y. Xu, X. Ding, Y. Zheng, R. Nishinakamura, G.L. Xu, H. Wang, S. Chen, S. Gao and B. Zhu, Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of 5-Methylcytosine, Mol Cell 64, 913-925 (2016)

17 K. Misawa, D. Mochizuki, A. Imai, Y. Misawa, S. Endo, M. Mima, H. Kawasaki, T.E. Carey and T. Kanazawa, Epigenetic silencing of SALL3 is an independent predictor of poor survival in head and neck cancer, Clinical epigenetics 9, 64 (2017)

18 K. Misawa, Y. Misawa, A. Imai, D. Mochizuki, S. Endo, M. Mima, R. Ishikawa, H. Kawasaki, T. Yamatodani and T. Kanazawa, Epigenetic modification of SALL1 as a novel biomarker for the prognosis of early stage head and neck cancer, J Cancer 9, 941-949 (2018)

19 A. Imai, D. Mochizuki, Y. Misawa, T. Nakagawa, S. Endo, M. Mima, S. Yamada, H. Kawasaki, T. Kanazawa and K. Misawa, SALL2 Is a Novel Prognostic Methylation Marker in Patients with Oral Squamous Carcinomas: Associations with SALL1 and SALL3 Methylation Status, DNA Cell Biol 38, 678-687 (2019)

20 D. Mochizuki, Y. Misawa, H. Kawasaki, A. Imai, S. Endo, M. Mima, S. Yamada, T. Nakagawa, T. Kanazawa and K. Misawa, Aberrant Epigenetic Regulation in Head and Neck Cancer Due to Distinct EZH2 Overexpression and DNA Hypermethylation, Int J Mol Sci 19, (2018)

21 K. Misawa, D. Mochizuki, A. Imai, S. Endo, M. Mima, Y. Misawa, T. Kanazawa, T.E. Carey and H. Mineta, Prognostic value of aberrant promoter hypermethylation of tumor-related genes in early-stage head and neck cancer, Oncotarget 7, 26087-26098 (2016)

22 W.Y. Huang, S.D. Hsu, H.Y. Huang, Y.M. Sun, C.H. Chou, S.L. Weng and H.D. Huang, MethHC: a database of DNA methylation and gene expression in human cancer, Nucleic Acids Res 43, D856-861 (2015)

23 J. Itou, W. Li, S. Ito, S. Tanaka, Y. Matsumoto, F. Sato and M. Toi, Sal-like 4 protein levels in breast cancer cells are post-translationally down-regulated by tripartite motif-containing 21, J Biol Chem 293, 6556-6564 (2018)

24 D. Cao, P.A. Humphrey and R.W. Allan, SALL4 is a novel sensitive and specific marker for metastatic germ cell tumors, with particular utility in detection of metastatic yolk sac tumors, Cancer 115, 2640-2651 (2009)

25 C. Sun, P. Lan, Q. Han, M. Huang, Z. Zhang, G. Xu, J. Song, J. Wang, H. Wei, J. Zhang, R. Sun, C. Zhang and Z. Tian, Oncofetal gene SALL4 reactivation by hepatitis B virus counteracts miR-200c in PD-L1-induced T cell exhaustion, Nat Commun 9, 1241 (2018)

26 W. Du, L. Ni, B. Liu, Y. Wei, Y. Lv, S. Qiang, J. Dong and X. Liu, Upregulation of SALL4 by EGFR activation regulates the stemness of CD44-positive lung cancer, Oncogenesis 7, 36 (2018)

27 J.U. Marquardt and S.S. Thorgeirsson, Sall4 in "stemness"-driven hepatocarcinogenesis, N Engl J Med 368, 2316-2318 (2013)

28 S.S. Zeng, T. Yamashita, M. Kondo, K. Nio, T. Hayashi, Y. Hara, Y. Nomura, M. Yoshida, T. Hayashi, N. Oishi, H. Ikeda, M. Honda and S. Kaneko, The transcription factor SALL4 regulates stemness of EpCAM-positive hepatocellular carcinoma, J Hepatol 60, 127-134 (2014)

29 L. Vilorio-Marques, V. Martin, C. Diez-Tascon, M.F. Gonzalez-Sevilla, T. Fernandez-Villa, E. Honrado, V. Davila-Batista and A.J. Molina, The role of EZH2 in overall survival of colorectal cancer: a meta-analysis, Scientific reports 7, 13806 (2017)

30 J. Itou, Y. Matsumoto, K. Yoshikawa and M. Toi, Sal-like 4 (SALL4) suppresses CDH1 expression and maintains cell dispersion in basal-like breast cancer, FEBS Lett 587, 3115-3121 (2013)

31 J. He, M. Zhou, X. Chen, D. Yue, L. Yang, G. Qin, Z. Zhang, Q. Gao, D. Wang, C. Zhang, L. Huang, L. Wang, B. Zhang, J. Yu and Y. Zhang, Inhibition of SALL4 reduces tumorigenicity involving epithelial-mesenchymal transition via Wnt/beta-catenin pathway in esophageal squamous cell carcinoma, J Exp Clin Cancer Res 35, 98 (2016)

32 K. Nimura, K. Ura, H. Shiratori, M. Ikawa, M. Okabe, R.J. Schwartz and Y. Kaneda, A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf-Hirschhorn syndrome, Nature 460, 287-291 (2009)

33 J. Yang, SALL4 as a transcriptional and epigenetic regulator in normal and leukemic hematopoiesis, Biomark Res 6, 1 (2018)

34 T. Baubec, D.F. Colombo, C. Wirbelauer, J. Schmidt, L. Burger, A.R. Krebs, A. Akalin and D. Schubeler, Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation, Nature 520, 243-247 (2015)

35 J. Lu, H.W. Jeong, N. Kong, Y. Yang, J. Carroll, H.R. Luo, L.E. Silberstein, Yupoma and L. Chai, Stem cell factor SALL4 represses the transcriptions of PTEN and SALL1 through an epigenetic repressor complex, PLoS One 4, e5577 (2009)

36 A. Li, Y. Jiao, K.J. Yong, F. Wang, C. Gao, B. Yan, S. Srivastava, G.S. Lim, P. Tang, H. Yang, D.G. Tenen and L. Chai, SALL4 is a new target in endometrial cancer, Oncogene 34, 63-72 (2015)

37 X. Zhang, X. Yuan, W. Zhu, H. Qian and W. Xu, SALL4: an emerging cancer biomarker and target, Cancer Lett 357, 55-62 (2015)

38 Y. Ma, W. Cui, J. Yang, J. Qu, C. Di, H.M. Amin, R. Lai, J. Ritz, D.S. Krause and L. Chai, SALL4, a novel oncogene, is constitutively expressed in human acute myeloid leukemia (AML) and induces AML in transgenic mice, Blood 108, 2726-2735 (2006)

39 X. Yue, L. Xiao, Y. Yang, W. Liu, K. Zhang, G. Shi, H. Zhou, J. Geng, X. Ning, J. Wu and Q. Zhang, High cytoplasmic expression of SALL4 predicts a malignant phenotype and poor prognosis of breast invasive ductal carcinoma, Neoplasma 62, 980-988 (2015)

40 L. Zhang, Y. Yan, Y. Jiang, Y. Cui, Y. Zou, J. Qian, C. Luo, Y. Lu and X. Wu, The expression of SALL4 in patients with gliomas: high level of SALL4 expression is correlated with poor outcome, J Neurooncol 121, 261-268 (2015)

41 S. Onder, O.C. Taskin, F. Sen, S. Topuz, S. Kucucuk, H. Sozen, R. Ilhan, S. Tuzlali and E. Yavuz, High expression of SALL4 and fascin, and loss of E-cadherin expression in undifferentiated/dedifferentiated carcinomas of the endometrium: An immunohistochemical and clinicopathologic study, Medicine (Baltimore) 96, e6248 (2017)

42 N. Yanagihara, D. Kobayashi, K. Kuribayashi, M. Tanaka, T. Hasegawa and N. Watanabe, Significance of SALL4 as a drugresistant factor in lung cancer, Int J Oncol 46, 1527-1534 (2015)

43 M. Miettinen, Z. Wang, P.A. McCue, M. Sarlomo-Rikala, J. Rys, W. Biernat, J. Lasota and Y.S. Lee, SALL4 expression in germ cell and non-germ cell tumors: a systematic immunohistochemical study of 3215 cases, Am J Surg Pathol 38, 410-420 (2014)

44 S. Zhou, R. Venkatramani, E. Gomulia, N. Shillingford and L. Wang, The diagnostic and prognostic value of SALL4 in hepatoblastoma, Histopathology 69, 822-830 (2016)

45 X. Zhang, P. Zhang, M. Shao, X. Zang, J. Zhang, F. Mao, H. Qian and W. Xu, SALL4 activates TGF-beta/SMAD signaling pathway to induce EMT and promote gastric cancer metastasis, Cancer Manag Res 10, 4459-4470 (2018)

Figure legends

Figure 1. *SALL4* mRNA patterns in matched pairs of HNSCC tissue and adjacent normal mucosal tissues. Relative *SALL4* mRNA expression levels in (**A**) the test cohort and (**B**) the validation cohort assessed by qRT-PCR. Differences between cancerous and normal mucosal tissues were considered significant, as determined by Student's *t*-test. (**C**) The area under the ROC curve (AUROC) value for *SALL4* was 0.771. At a cut-off value of 7.438, the sensitivity was 79.87% and the specificity 65.77%, respectively.

Figure 2. Effects of siRNA-mediated SALL4 knockdown in FaDu and UM-SCC-6 cells. (**A**) Western blot analysis of SALL4 protein expression after siRNA-mediated *SALL4* knockdown. (**B**) Chart illustrating the quantification of *SALL4* mRNA expression levels assessed by qRT-PCR after siRNA-mediated *SALL4* knockdown. (**C**) Colony formation upon transfection with SALL4 siRNA or si-NC. (**D**) *DNMT3A* mRNA expression assessed by qRT-PCR after siRNA-mediated *SALL4* knockdown. (**E**) Effect of *SALL4* knockdown on expression of the *DNMT3B* gene. The experiments were performed in duplicate and repeated thrice. The mean ± standard deviation value for each treatment is shown. NC: negative control, $*p < 0.05$.

Figure 3. Comparison of methylation rates in 12 tumour suppressor genes with *SALL4* expression in primary HNSCC tissues. (**A**) Distribution of *SALL4* gene expression status and promoter methylation of 12 tumour suppressor genes. Shaded boxes indicate the presence of methylation, and open boxes indicate the absence of methylation. (**B**) Correlation between gene methylation rate and *SALL4* expression status in HNSCC patients ($p = 0.001$). The gene methylation rates for the different groups were compared using Student's *t*-test. (**C**) Relationship between *SALL4* and *DNMT3A* mRNA levels in 161 HNSCC tissues ($p < 0.001$).

(**D**) *DNMT3B* expression was not associated with *SALL4* expression status ($p = 0.397$). ** $p < 0.01$.

Figure 4. Kaplan-Meier survival curves of HNSCC patients based on *SALL4* gene expression status. DFS for (**A**) the test cohort (n = 110), (**B**) the validation cohort (n = 120) and (**C**) the test and validation cohorts combined (n = 230). (**D**) tumour size in T1 and T2 cases (n = 105), (**E**) lymph node status in N0 cases (n = 106) and in (**F**) stages I and II cases (n = 61). Gray and black lines indicate patients with tumours showing *SALL4* downregulation and upregulation, respectively. * $p < 0.05$.

Supplementary Fig. S1.

Kaplan–Meier survival curves for patients with (**A**) oral cancer (n = 76; $p = 0.026$), (**B**) hypopharyngeal cancer (n = 54; $p = 0.420$), (**C**) laryngeal cancer (n = 51; $p = 0.173$), and (**D**) oropharyngeal cancer (n = 35; $p = 0.111$). * $p < 0.05$.

Supplementary Fig. S2.

Data regarding *SALL4* (**A**), *DNMT3A* (**B**)**,** and *DNMT3B* (**C**) mRNA expression in head and neck squamous cell carcinoma were obtained from TCGA (https://tcga-data.nci.nih.gov/tcga/) and MethHC (http://methhc.mbc.nctu.edu.tw/php/index.php). * $p < 0.05$, ** $p < 0.01$.

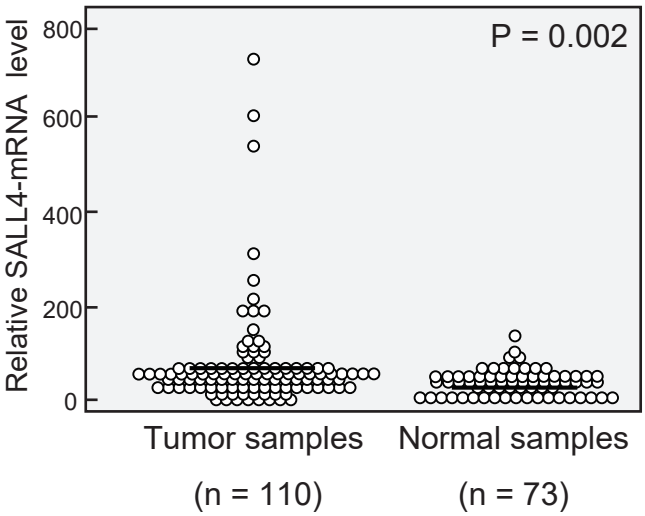
Supplementary Fig. S3.

Data regarding *SALL4*, *CDH1* (**A**), *COL1A2* (**B**) and *CDH13* (**C**) mRNA expression in head and neck squamous cell carcinoma were obtained from TCGA (https://tcga-

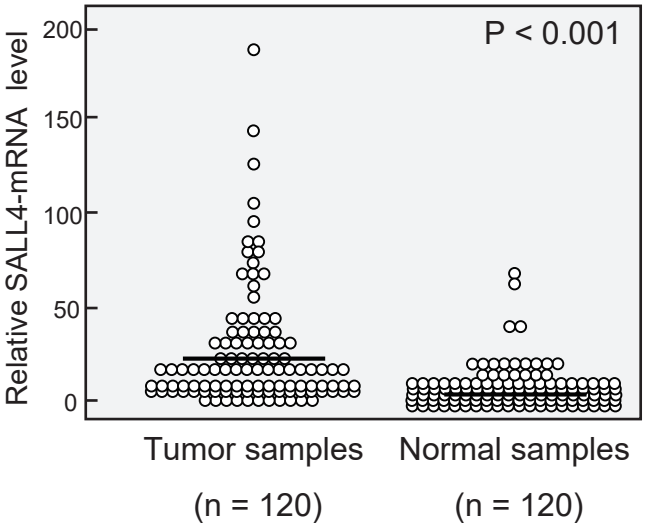
data.nci.nih.gov/tcga/) and MethHC (http://methhc.mbc.nctu.edu.tw/php/index.php). * $p <$ 0.05.

Figure 1 Kiyoshi Misawa

A Original cohort



B Validation cohort



C

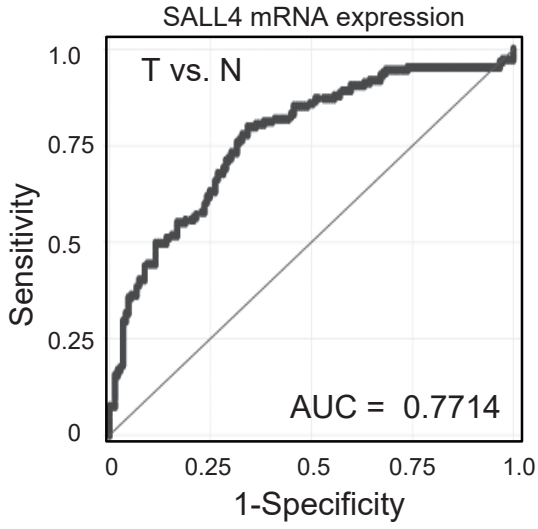


Figure 2 Kiyoshi Misawa

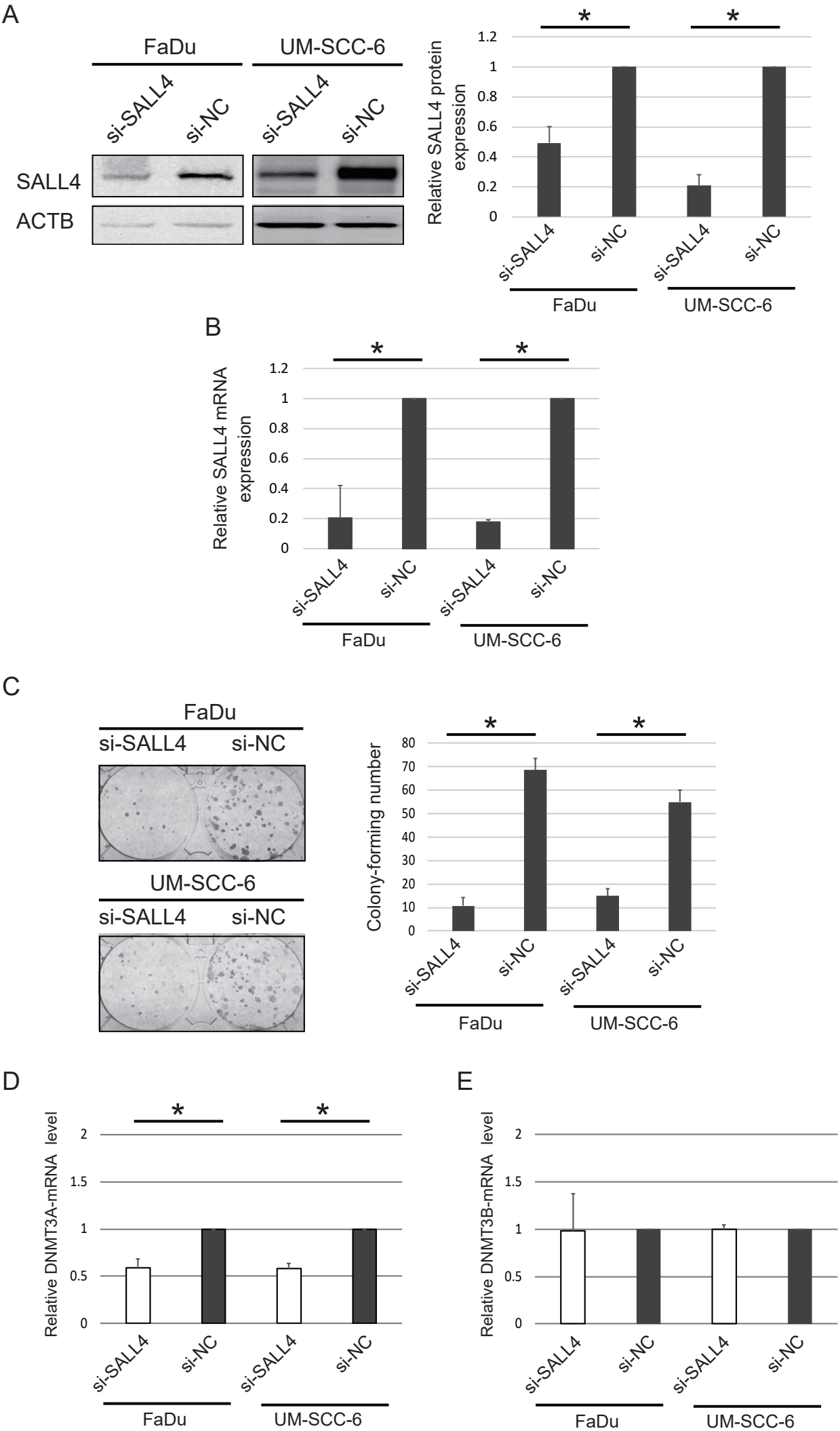
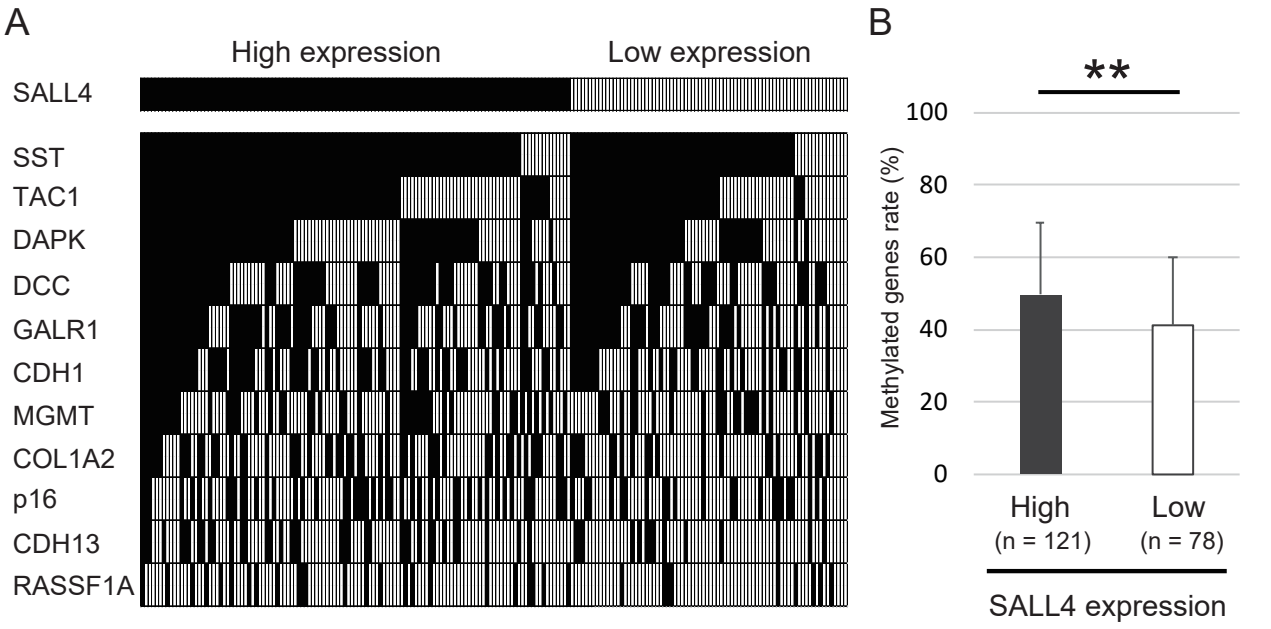
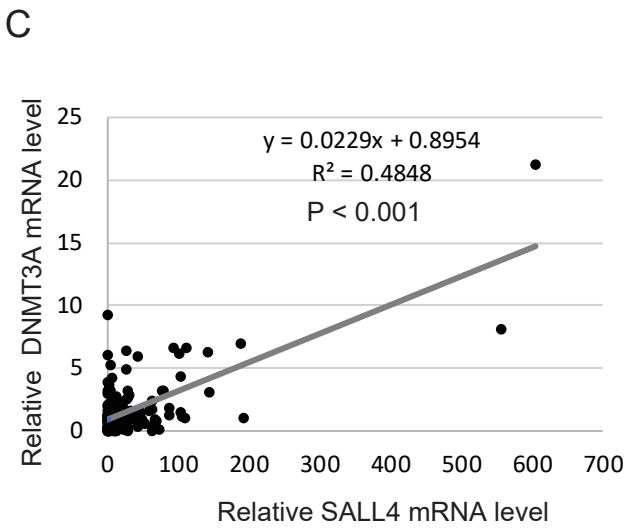


Figure 3 Kiyoshi Misawa





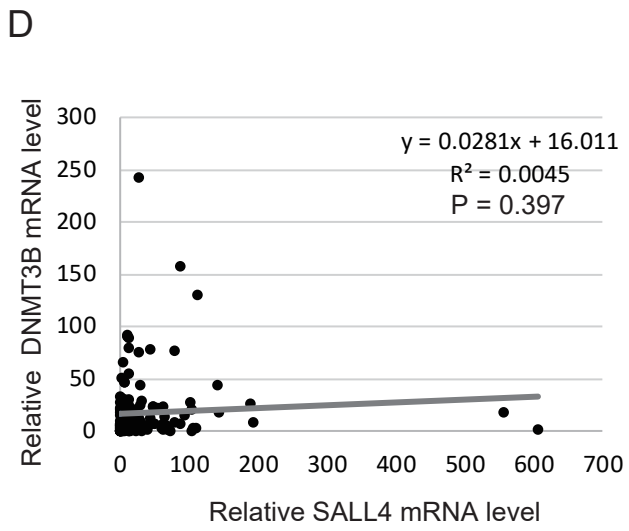
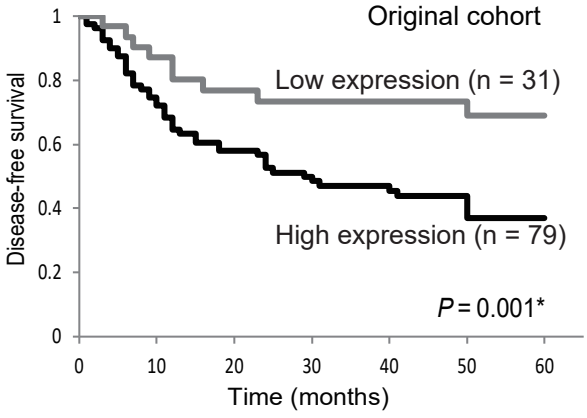


Figure 4 Kiyoshi Misawa

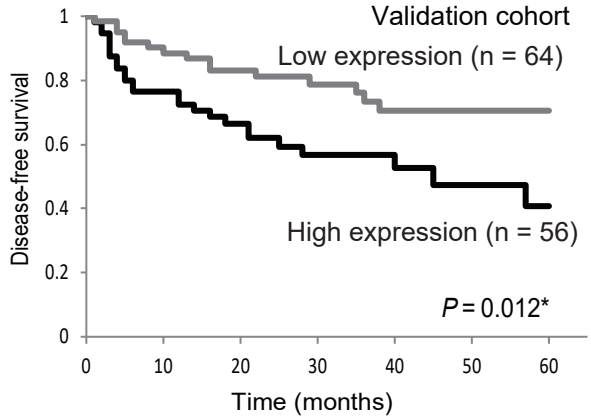
A



Number at risk

Time (Months)	0	10	20	30	40	50	60
Low expression	31	27	22	19	16	16	14
High expression	79	59	43	36	31	25	21

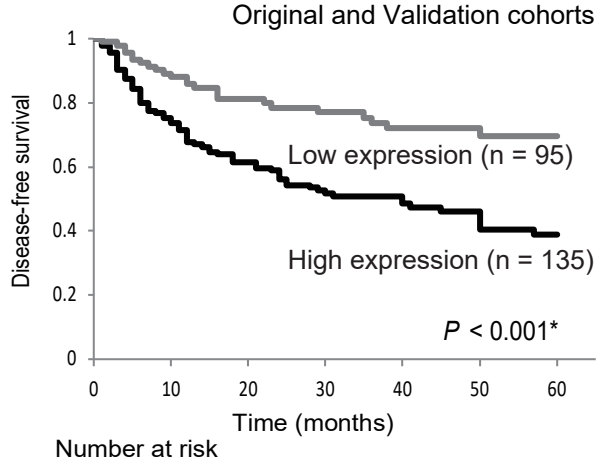
B



Number at risk

Time (Months)	0	10	20	30	40	50	60
Low expression	64	53	44	31	18	13	5
High expression	57	41	31	22	14	8	6

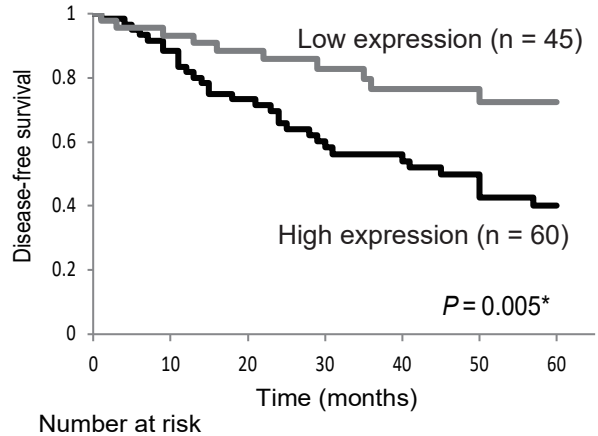
C



Number at risk

Time (Months)	0	10	20	30	40	50	60
Low expression	95	80	66	50	34	29	19
High expression	135	100	74	58	45	33	27

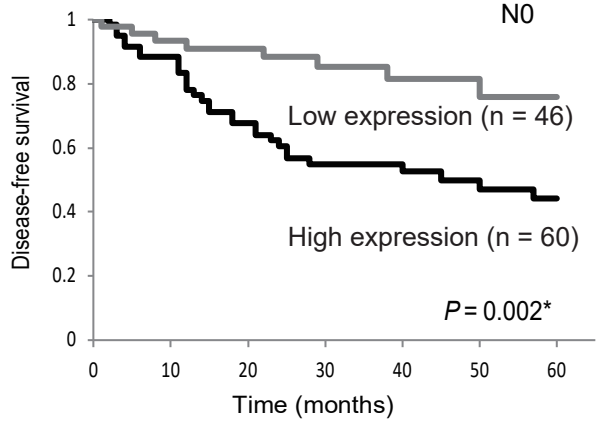
D



Number at risk

Time (Months)	0	10	20	30	40	50	60
Low expression	45	40	36	27	21	18	12
High expression	60	53	42	31	26	21	16

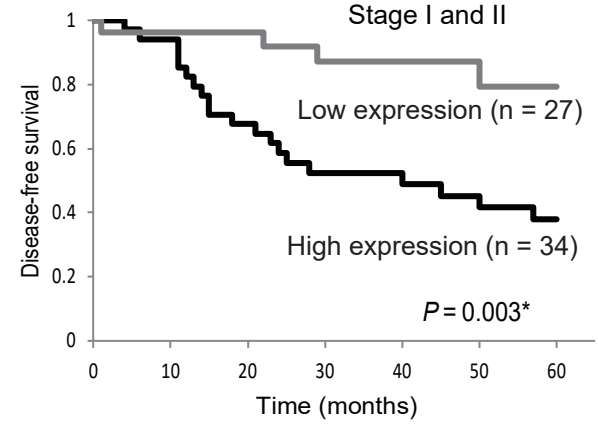
E



Number at risk

Time (Months)	0	10	20	30	40	50	60
Low expression	46	40	36	27	18	15	9
High expression	60	53	38	29	24	18	15

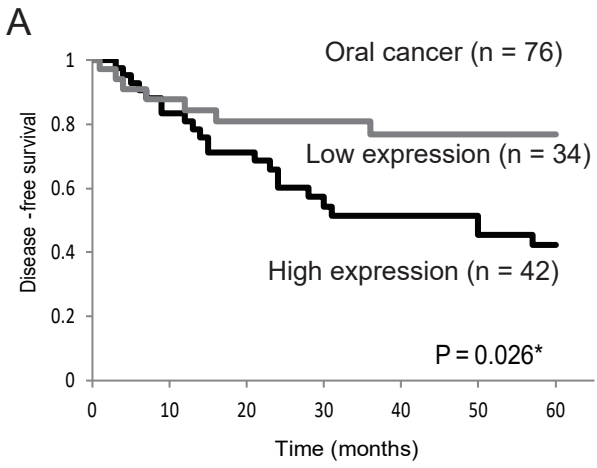
F

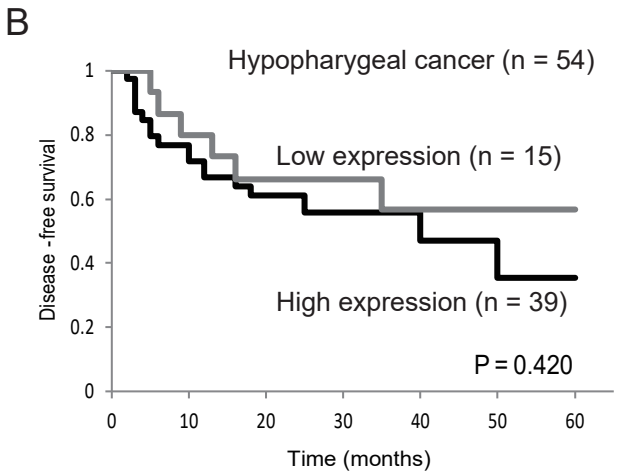


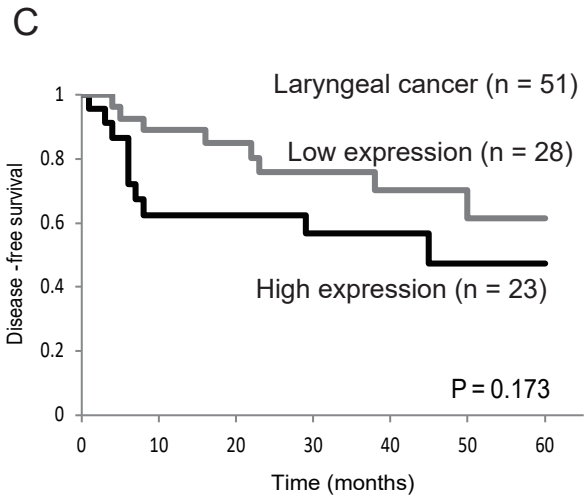
Number at risk

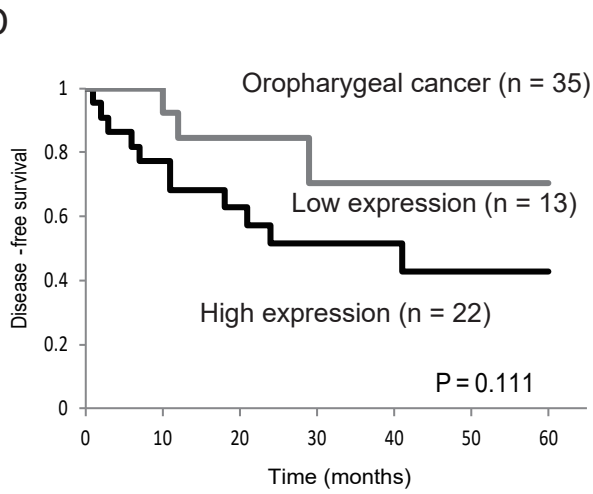
Time (Months)	0	10	20	30	40	50	60
Low expression	27	24	23	17	14	11	8
High expression	34	32	23	16	15	13	10

Supplementary Figure S1 Misawa K

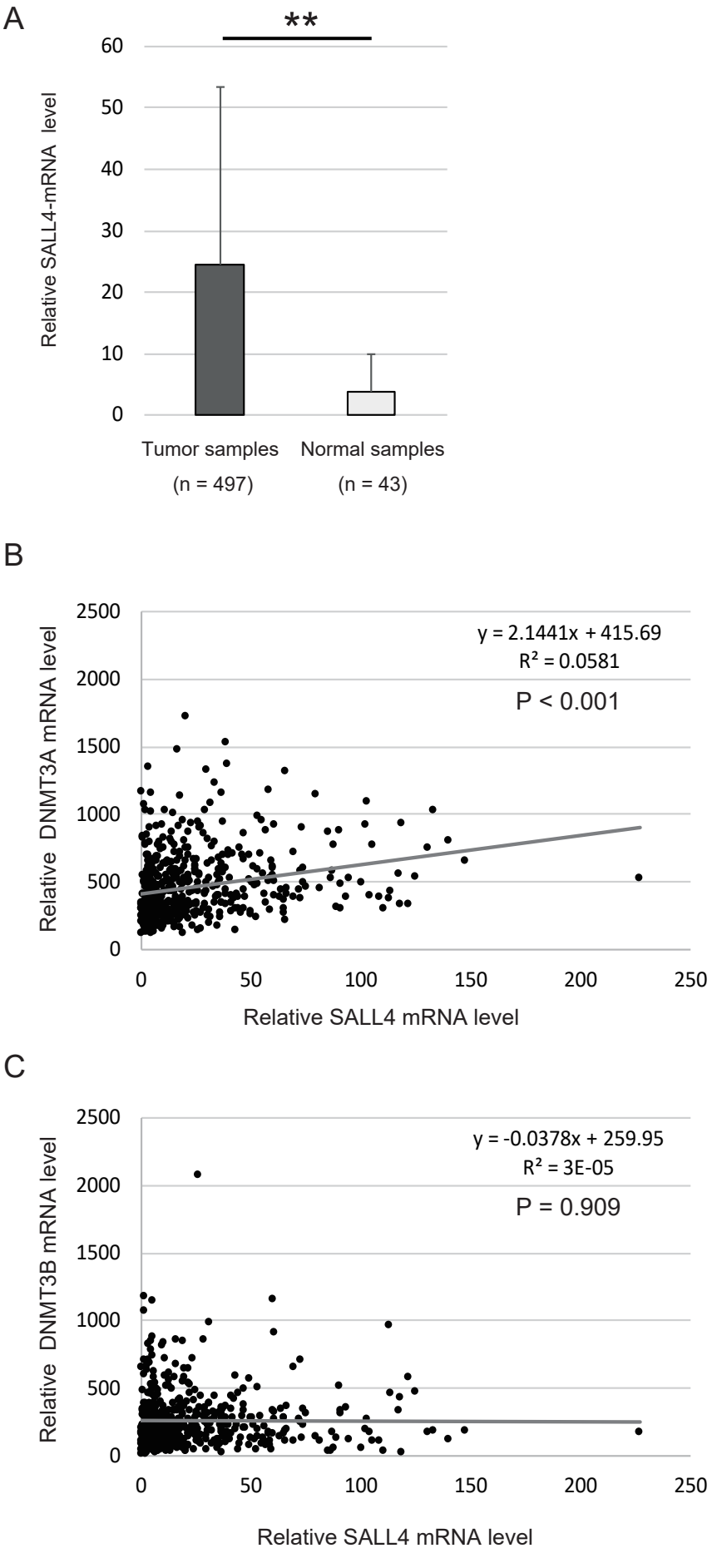






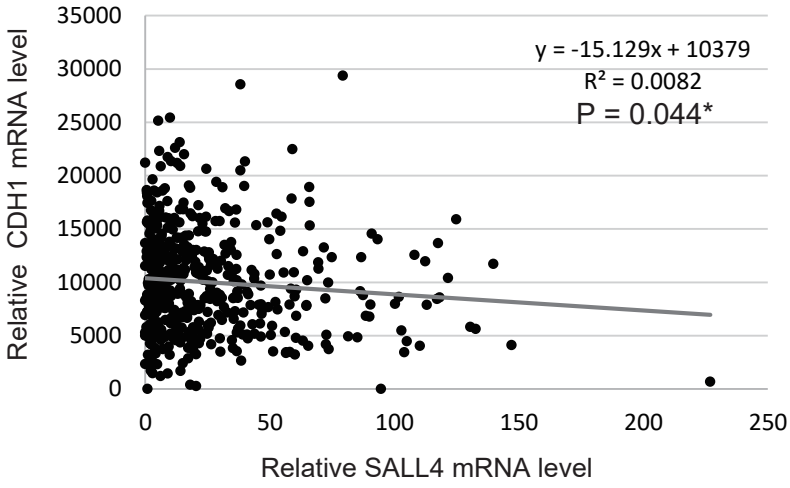


Supplementary Figure S2 Misawa K

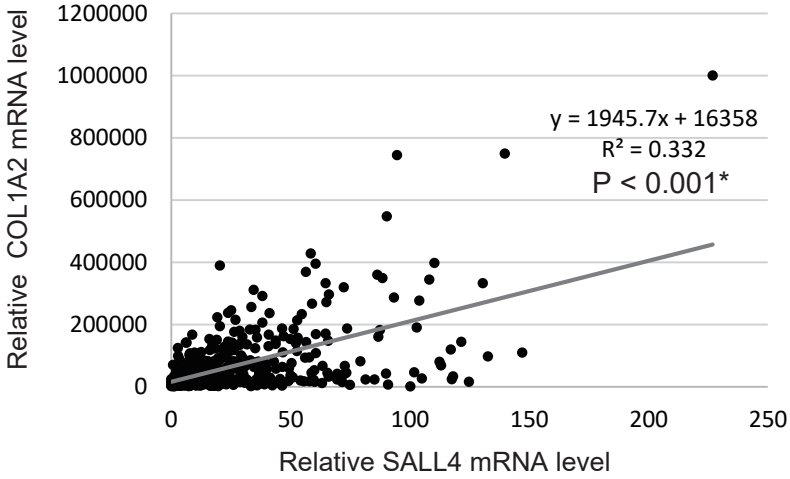


Supplementary Figure S2 Misawa K

A



B



C

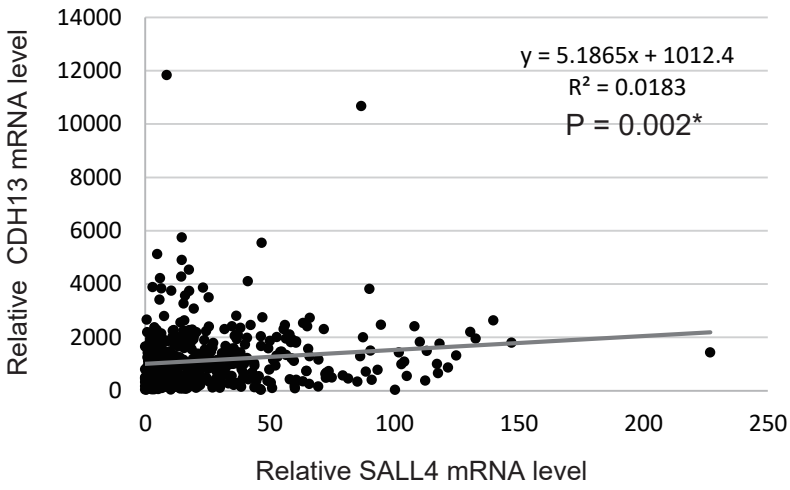


Table 1. *SALL4* Gene Expression Status in HNSCC Primary Samples. †Fisher's exact probability test. * P < 0.05

	Original cohort (n = 110)			Validation cohort (n = 120)			Original and Validation cohorts (n = 230)		
Patient and tumor characteristics	High (79)	Low (31)	*P*-value†	High (56)	Low (64)	*P*-value†	High (135)	Low (95)	*P*-value†
Age									
Under 65	40	17		21	25		61	42	
65 and older	39	14	0.832	35	39	1	74	53	0.894
Gender									
Male	62	25		52	51		114	76	
Female	17	6	1	4	13	0.064	21	19	1
Alcohol exposure									
Ever	42	18		50	49		92	67	
Never	37	13	0.676	6	15	0.092	43	28	0.772
Smoking status									
Smoker	51	21		50	43		101	64	
Non-smoker	28	10	0.826	6	21	0.004*	34	31	1
Tumor size									
T1-2	40	17		20	29		60	46	
T3-4	39	14	0.832	36	35	0.353	75	49	1
Lympho-node status									
N0	35	16		24	31		59	47	
N+	44	15	0.529	32	33	0.585	76	48	1
Stage									
I, II, III	36	16		25	27		61	43	
IV	43	15	0.672	31	37	1	74	52	1
Recurrence events									
Positive	47	9		25	15		72	24	
Negative	32	22	0.006*	31	49	0.020*	63	71	< 0.001*

Table 2. Multivariate analysis of factors affecting survival using Cox proportional hazards model in 230 HNSCC patients.

	Disease-free survival		
Variables	HR	(95% CI)	*P*
Age			
65 and older vs. < 65	1.097	(0.730-1.647)	0.656
Sex			
Male vs. Female	1.024	(0.598-1.754)	0.930
Alcohol exposure			
Ever vs. Never	0.529	(0.316-0.884)	0.015*
Smoking status			
Smoker ve. Non smoker	1.539	(0.882-2.688)	0.129
Stage			
I, II, III vs. IV	1.768	(1.152-2.714)	0.009*
SALL4 high expression			
Yes vs. No	2.566	(1.598-4.121)	< 0.0001*

HR: hazard ratio
95% CI: 95% confidence interval

Supplemental Table S1. Baseline Characteristics of the 230 Patients.

Patient and tumor characteristics	Original cohort (n = 110)	Validation cohort (n = 120)	Original and Validation cohorts (n = 230)
Age			
Under 65	57 (51.8%)	46 (38.3%)	103 (44.8%)
65 and older	53 (48.2%)	74 (61.7%)	127 (55.2%)
Gender			
Male	87 (79.1%)	103 (85.8%)	190 (82.6%)
Female	23 (20.9%)	17 (14.2%)	40 (17.4%)
Alcohol exposure			
Ever	60 (54.5%)	99 (82.5%)	159 (69.1%)
Never	50 (45.5%)	21 (17.5%)	71 (30.9%)
Smoking status			
Smoker	72 (65.5%)	93 (77.5%)	165 (71.7%)
Non-smoker	38 (34.5%)	27 (22.5%)	65 (28.3%)
Tumor size			
T1	11 (10.0%)	12 (10.0%)	23 (10.0%)
T2	46 (41.8%)	37 (30.8%)	83 (36.1%)
T3	21 (19.1%)	26 (21.7%)	47 (20.4%)
T4	32 (29.1%)	45 (37.5%)	77 (33.5%)
Lympho-node status			
N0	51 (46.4%)	55 (45.8%)	106 (46.1%)
N+	59 (53.6%)	65 (54.2%)	124 (53.9%)
Stage			
I	7 (6.4%)	10 (8.3%)	17 (7.4%)
II	26 (23.6%)	18 (15.0%)	44 (19.1%)
III	19 (17.3%)	24 (20.0%)	43 (18.7%)
IV	58 (52.7%)	68 (56.7%)	126 (54.8%)
Recurrence events			
Positive	56 (50.9%)	40 (33.3%)	96 (41.7%)
Negative	54 (49.1%)	80 (66.7%)	134 (58.3%)

Table S2. Q-RT-PCR and Q-MSP primer list

PCR	Gene	Forward primer 5'-3'	Reverse primer 5'-3'
Q-RT-PCR	SALL4	AGTATCAGAGCCGAAGCCCAGA	GGGCTCGGATAAACGTGGAA
Q-RT-PCR	DNMT3A	AGTACGACGACGACGGCTA	CACACTCCACGCAAAAGCAC
Q-RT-PCR	DNMT3B	AGGGAAGACTCGATCCTCGTC	GTGTGTAGCTTAGCAGACTGG
Q-RT-PCR	GAPDH	GCACCGTCAAGGCTGAGAAC	TGGTGAAGACGCCAGTCTCTA
Q-MSP	SST	GGGGCGTTTTTTAGTTTGACGT	AACAACGATAACTCCGAACCTCG
Q-MSP	TAC1	GGCGGTTAATTAAATATTGAGCAGAAAGTCGC	AAATCCGAACGCGCTCTTTCG
Q-MSP	DAPK	GGATAGTCGGATCGAGTTAACGTC	CCCTCCCAAACGCCGA
Q-MSP	DCC	TTGTTCGCGATTTTTGGTTTC	ACCGATTACTTAAAAATACGCG
Q-MSP	GALR1	GGTTCGCGGTATTCGGTAGT	GGTTCGCGGTATTCGGTAGT
Q-MSP	CDH1	GTGGGCGGGTCGTTAGTTTC	ACCACAACCAATCAACGCGA
Q-MSP	MGMT	TTCGACGTTCGTAGGTTTTCGC	GCACTCTTCCGAAAACGAAACG
Q-MSP	COL1A2	ACGGTAGTAGGAGGTTTCGG	CGCAAAACCCCTAAATCACCGACG
Q-MSP	p16	GTATTTTTTTCGAGTATTCGTTTACGGC	CAAATCCTCTAAAAAAACCGCGA
Q-MSP	CDH13	TTTGGGAAGTTGGTTGGTTGGC	ACTAAAAACGCCCGACGACG
Q-MSP	RASSF1A	CGTTCGGTTCGCGTTTGTTAGC	TAACCCGATTAAACCCGTACTTCG
Q-MSP	SALL1	GTCGTCGTTCGATTTTCGTAA	CGCTTACTTCCTCCGCGACA
Q-MSP	SALL2	CGGGAATGTTTCGGCGAAAG	CAAACGCGCTAAAACCTTCGCA
Q-MSP	SALL3	GGGGTTCGAGCGTCGTTAGT	CCGTACTCGAAAACCCCGTC
Q-MSP	ACTB	TGGTGATGGAGGAGGTTTAGTAAGT	AACCAATAAAACCTACTCCTCCCTTAA

Table S3. SALL4 expression status with the methylation of other 11 genes.

Genes	Methylation status	SALL4 high (N = 121)	SALL4 low (N = 78)	P-values†
SST	Methylated	107	63	
	Unmethylated	14	15	0.053
TAC1	Methylated	81	45	
	Unmethylated	40	33	1
DAPK	Methylated	69	46	
	Unmethylated	52	32	0.883
DCC	Methylated	69	41	
	Unmethylated	52	37	1
GALR1	Methylated	63	40	
	Unmethylated	58	38	1
CDH1	Methylated	67	28	
	Unmethylated	54	50	0.009*
MGMT	Methylated	48	23	
	Unmethylated	73	55	0.173
COL1A2	Methylated	50	21	
	Unmethylated	71	57	0.048*
P16	Methylated	47	22	
	Unmethylated	74	56	0.131
CDH13	Methylated	37	14	
	Unmethylated	84	64	0.048*
RASSF1A	Methylated	24	9	
	Unmethylated	97	69	0.171

† Fisher's exact probability test.

* $P < 0.05$.

Table S4. SALL4 expression status with the methylation of other *SALL* genes.

Genes	Methylation status	SALL4 high (N = 112)	SALL4 low (N = 83)	P-values†
SALL1	Methylated	42	38	
	Unmethylated	70	45	1
SALL2	Methylated	31	24	
	Unmethylated	81	59	1
SALL3	Methylated	75	46	
	Unmethylated	37	37	0.104

† Fisher's exact probability test.

* P <0.05.

Supplementary Table S5 Correlation between primary tumor sites and *SALL4* gene expression status.

Primary site		Oral cavity			Hypopharynx			Larynx			Oropharynx			Paranasal cavity		
Expression status		High	Low		High	Low		High	Low		High	Low		High	Low	
Characteristics	Overall (%)	42 (55.3%)	34 (44.7%)	P †	39 (72.2%)	15 (27.8%)	P †	23 (45.1%)	28 (54.9%)	P †	22 (62.9%)	13 (37.1%)	P †	9 (64.3%)	5 (35.7%)	P †
Age	< 65	23	15		13	8		7	11		11	7		5	3	
	> 65	19	19	0.369	26	7	0.220	16	17	0.567	11	6	1	4	2	1
Gender	Male	33	23		33	14		23	26		19	10		6	3	
	Female	9	11	0.307	6	1	0.659	0	2	0.560	3	3	0.801	3	2	1
Alcohol exposure	drinker	30	15		31	12		17	24		15	9		5	1	
	non drinker	12	19	0.020*	8	3	1	6	4	0.483	7	4	1	4	4	0.301
Smoking status	smoker	29	21		29	13		20	17		16	10		7	0	
	non smoker	13	13	0.024*	10	2	0.474	3	11	0.058	6	3	1	2	5	0.026*
Tumor size	T1-2	29	21		13	7		5	9		12	6		1	2	
	T3-4	13	13	0.628	26	8	1	18	19	0.533	10	7	0.733	8	3	0.560
Lympho-node status	N0	19	21		15	1		11	14		9	5		6	5	
	N+	23	13	0.172	24	14	0.024*	12	14	1	13	8	1	3	0	0.437
Stage	I, II, III	21	21		18	4		10	11		9	4		3	3	
	IV	21	13	0.358	21	11	0.230	13	17	1	13	9	0.721	6	2	0.687
Recurrence events	positive	22	7		21	6		10	8		11	3		8	0	
	negative	20	27	0.008*	18	9	0.384	13	20	0.378	11	10	0.162	1	5	0.008*

† Fisher's exact test * P<0.05