Hypermethylation of collagen α2 (I) gene (COL1A2) is an independent predictor of survival in head and neck cancer

Kiyoshi Misawa\textsuperscript{a},\textsuperscript{*}, Takeharu Kanazawa\textsuperscript{b}, Yuki Misawa\textsuperscript{a}, Atsushi Imai\textsuperscript{a}, Shiori Endo\textsuperscript{a}, Katsura Hakamada\textsuperscript{a} and Hiroyuki Mineta\textsuperscript{a}  
\textsuperscript{a}Department of Otolaryngology/Head and Neck Surgery, Hamamatsu University School of Medicine, Shizuoka, Japan  
\textsuperscript{b}Department of Otolaryngology/Head and Neck Surgery, Jichi Medical University, Tochigi, Japan

Abstract. Objectives: Collagen production plays a role in the development of tumors from cancer cells. The aim of the present study is to examine the involvement of epigenetic alteration of Collagen α2 (I) (COL1A2) gene expression in cases of head and neck squamous cell carcinoma (HNSCC).  
Methods: COL1A2 expression was examined in a panel of cell lines using RT-PCR. The methylation status of the COL1A2 promoter was studied using bisulfate sequencing and methylation-specific PCR (MSP).  
Results: COL1A2 expression was absent in 6 of 11 (54.5%) UM-SCC cell lines, whereas three nonmalignant cell lines had stable expressions. MSP analysis showed that 46/98 (46.9%) contained methylated alleles. COL1A2 methylation was significantly correlated with tumor size ($P = 0.041$), lymph node status ($P = 0.008$), tumor stage ($P = 0.011$), H-cadherin methylation ($P = 0.039$) and disease-free survival ($P = 0.005$). On multivariate Cox proportional hazard regression, which included age, sex, smoking status, and alcohol exposure, both tumor stage and COL1A2 methylation remained independent prognostic factors.  
Conclusions: This study suggests that CpG hypermethylation is a likely mechanism of COL1A2 gene inactivation, supporting the hypothesis that the COL1A2 gene may play a role in the tumorigenesis of HNSCC and may serve as an important biomarker.  

Keywords: COL1A2, DNA methylation, CpG island, head and neck cancer, 5-azacytidine

1. Introduction

Squamous cell carcinoma of the head and neck region (HNSCC) is the sixth most frequent cancer and affects ~500,000 patients per year worldwide [1]. Various factors are linked to the development of HNSCC; sex, tobacco smoking, alcohol consumption, and HPV infection are major risk factors for the disease [2]. The prognosis for patients with HNSCC is still poor today despite tremendous technical advances in surgical treat-
detected, detection of methylated genes is an attractive biomarker for risk assessment, early detection, and monitoring of prognosis, as well as the prevention of HNSCC.

Collagen is a large family of at least 28 extracellular matrix proteins that play vital structural and physiological roles in maintaining the integrity of and contributing to the homeostasis of the human body [9]. Type I collagen, the most abundant collagen molecule, normally consists of a heterotrimer of two α1 (I) (produced by the COL1A1) chains and one α2 (I) (produced by the COL1A2) chain [10]. Structural integrity and coordinated biosynthesis of the two chains are critically important for tissue morphogenesis, growth, homeostasis, and repair [11]. Hypermethylation of COL1A2 has been described in many cancers, mainly adenocarcinomas, including colorectal cancer [12], melanoma [13], bladder cancer [14], neuroblastoma [15], medulloblastoma [16,17], breast cancer [18], and hepatoma [19]. The investigation for oncogenic mechanism by COL1A2 gene inactivation was also recently progressing. Mori et al. reported that COL1A2 gene inactivation through CpG hypermethylation may contribute to proliferation and migration activity of bladder cancer [14]. Despite of accumulated knowledges in adenocarcinoma, hypermethylation of COL1A2 in squamous cell carcinoma (SCC) such as head and neck, esophagus, lung and cervix, is an area that still remains to be explored.

The aim of the present study is to evaluate COL1A2 methylation in large panels of primary HNSCC (that is typical SCC specimens). We demonstrate that loss of COL1A2 expression is associated with hypermethylation of key CpG sites within transcription factor binding domains and that expression can be restored after treatment of key CpG sites within transcription factor binding domains and that expression can be restored after treatment with the demethylating agent, 5-azacytidine, and the histone deacetylase inhibitor, trichostatin A (TSA). Structural integrity and coordinated biosynthesis of the two chains are critically important for tissue morphogenesis, growth, homeostasis, and repair [11]. Hypermethylation of COL1A2 has been described in many cancers, mainly adenocarcinomas, including colorectal cancer [12], melanoma [13], bladder cancer [14], neuroblastoma [15], medulloblastoma [16,17], breast cancer [18], and hepatoma [19]. The investigation for oncogenic mechanism by COL1A2 gene inactivation was also recently progressing. Mori et al. reported that COL1A2 gene inactivation through CpG hypermethylation may contribute to proliferation and migration activity of bladder cancer [14]. Despite of accumulated knowledges in adenocarcinoma, hypermethylation of COL1A2 in squamous cell carcinoma (SCC) such as head and neck, esophagus, lung and cervix, is an area that still remains to be explored.

The aim of the present study is to evaluate COL1A2 methylation in large panels of primary HNSCC (that is typical SCC specimens). We demonstrate that loss of COL1A2 expression is associated with hypermethylation of key CpG sites within transcription factor binding domains and that expression can be restored after treatment with the demethylating agent, 5-azacytidine, and the histone deacetylase inhibitor, trichostatin A (TSA). Furthermore, assessment of primary tumor specimens confirm that hypermethylation is as common in patient tumors and is directly associated with tumor size and metastasis. This study suggests that hypermethylation of COL1A2 gene in the primary tumor is an independent predictor of survival in head and neck cancer.

2. Materials and methods

2.1. Tumor samples and cell lines

Tumor specimens were obtained at surgery from 98 primary HNSCC samples. All patients were treated at the Department of Otolaryngology, Hamamatsu University School of Medicine. Clinical information including age, sex, tumor site, smoking status, alcohol intake, tumor size, lymph node status, and stage grouping were obtained from the clinical records. The average patient age was 63.6 years (range, 39–93 years), and the male: female ratio was 80:18. The primary tumor was located in the oral cavity (n = 39), hypopharynx (n = 25), larynx (n = 20), and oropharynx (n = 14). DNA and cDNA from 11 UM-SCC cell lines and 2 fibroblasts cell lines established from patients at the University of Michigan were kindly provided by Dr. Thomas E. Carey of the University of Michigan. Normal human keratinocytes (NHK) were a gift from Dr. No Hee Park of the UCLA School of Dentistry [20].

2.2. Bisulfite modification and MSP analysis

Bisulfite modification of genomic DNA was carried out as described previously [21]. Bisulfite-treated DNA was amplified by PCR with two designed pairs of methylation-specific PCR (MSP/UMSP) primers for the promoter region of the COL1A2 gene. The MSP primer set was 5’-ACGTTAGTAGGAGGTTTCGG-3’ (forward) and 5’-CGCAAACCCCCCTAAACACCCGA CG-3’ (reverse). The UMSP primer set was 5’-ATGGTAGTAGGAGGTTTCGG-3’ (forward) and 5’-CA CAAAACCCCCCTAAACACCCCAACA-3’ (reverse). The PCR conditions were: 94°C for 5 min; 39 cycles at 94°C for 30 s, 58°C (for detection of methylated DNA) or 54°C (for detection of unmethylated DNA) for 30 s, and 72°C for 40 s; and a final extension at 72°C for 5 min. The PCR products were 82 bp long. To analyze the methylation status of the p16 gene [5], E-cadherin gene [6], RASSF1A gene [7], H-cadherin gene [22], DAPK gene [23], MGMT gene [23], and DCC gene [24], primers and conditions as described previously were used. The PCR products were separated by electrophoresis through a 9% polyacrylamide gel and stained with ethidium bromide.

2.3. Bisulfite sequencing analysis for COL1A2

Bisulfite-treated DNA was amplified by bisulfite sequencing PCR (BSP) with primers that were specific for modified upper strand DNAs but did not contain any CpG sites. The BSP primer set was 5’-GTGTTTCTAATTTGGAAAGG-3’ (forward) and 5’-CTACAAAAACACAAATACTCC-3’ (reverse). The PCR conditions were: 94°C for 5 min; 45 cycles at 94°C for 30 s, 54°C for 30 s and 72°C for 40 s; and a final extension at 72°C for 5 min. The PCR products were 504 bp long. Bisulfite sequencing analysis was carried out as described previously [21].
2.4. RT-PCR for COL1A2

The primer sequences designed from the coding region of the human COL1A2 cDNA have been described previously [25]. The PCR products were separated by electrophoresis through a 9% polyacrylamide gel. Sense and antisense primers for the GAPDH gene used as an internal control have been described previously [21].

2.5. Reactivation of COL1A2 expression

Twelve hours after plating, cultures were incubated either for 48 h with 5-azacytidine (15 μg/ml, 30 μg/ml; Sigma, St. Louis, MO), an inhibitor of DNA methyltransferase, for 24 h with 300 nM trichostatin A (TSA) (Sigma), an inhibitor of histone deacetylase, or for 48 h with 5-azacytidine followed by 24 h incubation with TSA. The medium was then removed, and cultures were maintained in standard Dulbecco’s modified Eagle’s medium, which was replaced every other day [21].

2.6. Statistical analysis

For frequency analysis in contingency tables, statistical analyses of associations between variables were performed by Fisher’s exact test. The disease-free interval was measured from the date of the treatment to the date when locoregional recurrence or distant metastasis was diagnosed. Disease-free survival probabilities were estimated by the Kaplan-Meier method, and the log-rank test was applied to assess the significance of differences among actuarial survival curves. Cox proportional hazards multivariate regression analysis, which involved age, sex, smoking status, alcohol intake, stage grouping, and COL1A2 methylation, was used to identify the predictive value of the prognostic factors [6,26]. A significant difference was identified when the probability was less than 0.05. Statistical analysis was done with StatMate IV (ATMS Co., Ltd., Tokyo, Japan).

3. Results

3.1. COL1A2 expressions and effect of treatment with 5-azacytidine and TSA

A total of 11 UM-SCC cell lines were analyzed for COL1A2 expression using RT-PCR. COL1A2 expression was absent in 6 of 11 (54.5%) UM-SCC cell lines (UM-SCC-6, 10A, 10B, 14B, 22A, and 54) after 35 cycles (Fig. 1A). RT-PCR amplification of RNA from normal control cell lines (UM-81F, UM-99F and NHK) as template yielded a good signal after 35 cycles, indicating that COL1A2 is expressed in these normal cells (Fig. 1B). To establish that methylation was responsible for silencing COL1A2 gene expression, six cell lines (UM-SCC-6, 10A, 10B, 14B, 22A, and 54) were treated. These cell lines, which showed the absence of COL1A2 expression and hypermethylation by bisulfite sequencing analysis, were cultured with 5-azacytidine alone, TSA alone, or 5-azacytidine plus TSA. This facilitated assessment of the effect of
These primers were designed based on the bisulfite sequencing data. Forward primers for MSP and UMSP contain #15 and 16 CpG dinucleotides. Reverse primers for MSP and UMSP include #21-23 CpG dinucleotides (Fig. 2A). The MSP assay for COL1A2 correlated with the expression status. In all cell lines with reduced expression (UM-SCC-6, 10A, 10B, 14B, 22A, and 54), only methylated alleles were detected. In cell lines with apparently normal COL1A2 mRNA levels (UM-SCC-1, 2, 22B, 23, and 74B), both alleles were detected (Fig. 2B). Among 98 primary HNSCC DNA samples tested, the COL1A2 promoter was methylated in 46 (46.9%) cases and unmethylated in 52 (53.1%). Methylation of COL1A2 was significantly correlated with tumor size ($P = 0.041$), lymph node status ($P = 0.008$), and tumor stage ($P = 0.011$) (Fig. 2C, Table 1).

### 3.3. Bisulfite sequencing analysis

These cell lines were analyzed for COL1A2 promoter region methylation status. Bisulfite sequencing was used to more precisely map and determine the density of methylation throughout the COL1A2 promoter (Fig. 3A). CpG sites #1 to 20 were located upstream of TSS, and sites #21 to 25 were located downstream of TSS. We sequenced at least seven clones for each of the 11 cell lines. Cell lines UM-SCC-6, 10A, 10B, 14B, 22A, and 54 had loss of COL1A2 expression, as demonstrated by RT-PCR, and high methylation levels (62.9% to 90.5%). Cell lines UM-SCC-2, 22B, and 74B had stable COL1A2 expression by RT-PCR and moderate or low methylation levels, ranging from 38.3% to 59.0%. Methylation levels of UM-SCC-1 and 23 (68.0% and 78.0%) were higher than of UM-SCC-10B (62.9%), although these cell lines had positive COL1A2 expressions. For the six COL1A2-negative cell lines (UM-SCC-6, 10A, 10B, 14B, 22A, and 54), CpG site #15 upstream of the transcription start site was methylated in 91-100%. In contrast, for the cell lines with a more readily detectable message (UM-SCC-1, 2, 22B, 23, and 74B), the proportion of methylated alleles fell below 90%, and the message level of COL1A2 expression was stable (Fig. 3B).

### 3.4. Association of molecular markers and survival with COL1A2 methylation status

Methylation of COL1A2 was significantly correlated with $H$-cadherin methylation ($P = 0.039$). No other molecular markers showed a significant correlation with COL1A2 methylation status (Table 1).
univariate analysis, methylation of the COL1A2 gene was positively correlated with reduced disease-free survival \((P = 0.005, \text{log-rank test})\) (Fig. 4A). There was no significant difference in the disease-free survival of patients according to methylation patterns of the E-cadherin gene alone, the H-cadherin gene alone, or the combinations of the methylation changes of three genes (concordant pattern) (Fig. 4B, C, D). Moreover, results of multivariate Cox proportional hazard regression, which included age, sex, smoking status, alcohol exposure, and tumor stage, indicated that the group with methylation of the COL1A2 gene had a 2.104 times greater hazard than the group without methylation \((P = 0.031)\) (Table 2). Thus, COL1A2 promoter methylation is a predictor of poor outcome in patients with HNSCC.

4. Discussion

Differential DNA methylation at CpG islands has been associated with regulation of gene expression and is essential in normal development, X-chromosome inactivation, imprinting, suppression of parasitic DNA
K. Misawa et al. / Hypermethylation of COL1A2 in HNSCC

Fig. 3. Summary of bisulfite sequencing analysis. (A) Sequencing chromatograms is demonstrated CpG island methylation after bisulfite treatment in UM-SCC-14B. (B) The shading of each cell within the figure indicates the proportion of alleles that were found to be methylated by methylation-specific sequencing. The key is shown below each panel. The numbers in the top row indicate the CpG dinucleotide positions (labeled 1–25 in Fig. 1A) in the region amplified by the bisulfite sequencing analysis primers. The locations of the MSP primer binding sites (MSP/UMSP-F and MSP/UMSP-R) used for the data in Fig. 3B are indicated by the black boxes above the figure. The cell line numbers are shown in the left column. The numbers of clones sequenced (cl) are given in the second column. The column headed RT summarizes the RT-PCR results; RT(+) signifies positive for COL1A2 expression by RT-PCR; RT(−) signifies negative for COL1A2 expression.

sequences, and cancer [27]. Methylation has practical benefits over other molecular biomarkers. First, changes in methylation occur globally within a cancer’s DNA. Second, the abnormal result is a positive signal, reducing the problem of normal cell contamination [28]. There is increasing evidence showing that gene silencing due to aberrant methylation of DNA is an early event in carcinogenesis and is associated with early recurrence in some cancers [29]. The establishment of DNA methylation profiles of the primary tumor specimen might be a valuable tool in determining the prognosis and predicting the patient’s response to therapies [30].

It is interesting that the region surrounding the COL1A2 (7q22.1) transcription start site is not a typical CpG island. Therefore, this promoter may become methylated more easily than a protected CpG island promoter [12]. Sengupta et al. reported that the sequence-specific DNA-binding protein RFX (regulatory factor for X-box) could mediate the transcriptional activation of a methylated COL1A2 gene promoter that was repressed by methylation [31]. There is a binding site for the RFX family within the promoter region of the COL1A2 gene. Forward primers for MSP and UMSP bind in the segment of the promoter that contains CpG dinucleotide #15, which corresponds to a consensus RFX-binding site [12]. RFX1 was associated with distinct sets of co-repressors on the collagen transcription start site [10]. The DNA methyl-
Fig. 4. Kaplan-Meier Estimates of Disease-free Survival among 98 Case Patients. Data are reported for a three-gene panel consisting of COL1A2 (A), E-cadherin (B), and H-cadherin (C). Thirteen patients had methylation of all three genes, and 28 patients had no methylation of all three genes (D).

Transferase inhibitor, which inhibits DNA methylation, reduces RFX1/histone deacetylase 1 (HDAC1) binding to the collagen transcription start site. Trichostatin A stimulates the acetylation of RFX proteins and activates the collagen promoter [32]. In COL1A2 expression-silenced cell lines, their expression could be restored by demethylation with 5-azacytidine and TSA, inferring that DNA methylation and deacetylation of histones are major regulators of COL1A2 expression.

In this regard, it is worth noting that methylation of COL1A2 and cadherin genes may not be related directly, because, to the best of our knowledge, this is the first report to show a relation between the methylation statuses of these genes in various tumors, including HNSCC. Because the exact mechanisms of the functions of COL1A2 and cadherin genes in tumorigenesis are still unknown, it will be interesting to elucidate the relationships between the expressions of COL1A2 and cadherin genes. E-cadherin, encoded by CDH1, which is one of the best-characterized cadherins, is the key component for adherence junctions between epithelial cells [33]. E-cadherin gene promoter methylation is associated with increased pack years smoked in HNSCC [6]. E-cadherin hypermethylation is a prognostic marker of better survival and better radiotherapy response in HNSCC [34,35]. H-cadherin, encoded by CDH13, has also been found to be epigenetically silenced by promoter methylation in many types of tumors, including lung cancer, breast cancer, and colorectal cancer [36]. Recently, Broch et al. reported that methylation of p16 and H-cadherin was associated with early recurrence of stage I non-small cell lung cancer [29]. In nasopharyngeal carcinoma, H-cadherin was methylated in 89.7% of cases [37]. However, the roles of the H-cadherin gene have not been elucidated in other sites of primary HNSCC tumors. Methylation of H-cadherin was significantly correlated with DAPK methylation (P = 0.0009). (Data not shown)

There is also evidence for COL1A2 aberrant promoter methylation in genome-wide methylation stud-
ies of cancer cells, such as colorectal cancer [12], melanoma [13], bladder cancer [14], neuroblastoma [15], medulloblastoma [16,17], breast cancer [18], and hepatoma [19]. The investigation for oncogenic mechanism by COL1A2 gene inactivation was also recently progressing. In melanoma, COL1A2 methylation was predominantly detected in advanced stage tumors [13]. COL1A2 gene inactivation through CpG hypermethylation may contribute to proliferation and migration activity of bladder cancer [14]. The epigenetic status of COL1A2 may reflect the developmental biology of different medulloblastoma histological and molecular subtypes [16,17]. COL1A2 hypermethylation of CpG islands can be a promising marker of many types of tumors.

In this study, COL1A2 expression was shown to be frequently absent in HNSCC. The increased CpG dinucleotide #15 methylation of the COL1A2 gene is inversely correlated with COL1A2 mRNA levels. In a survey of 98 tumor tissue samples using MSP, hypermethylation of the COL1A2 promoter occurred with a high frequency (46.9%) (p16, 51.0%; E-cadherin, 39.8%; H-cadherin, 38.8%; DCC, 30.6%; MGMT, 28.6%; DAPK, 27.6%; and RASSF1A, 21.4%). Thus, COL1A2 resembled other major tumor suppressor genes in the frequency of aberrant promoter methylation. More importantly, in the case of COL1A2, there was a clear effect of methylation on clinical outcome. Disease-free survival was significantly worse in patients with methylation in COL1A2 (P = 0.005). On multivariate analysis, only COL1A2 methylation and clinical stage were significantly associated with poor survival when age, sex, smoking status, and alcohol intake were also considered. Thus, COL1A2 methylation appears important in disease progression and in the development of high-risk HNSCC.

In conclusion, the present study showed that the COL1A2 promoter methylation profile appears to be an important marker predicting the clinical outcome of HNSCC. This information can be used to identify patients with high-risk HNSCC who may benefit from adjuvant therapy and cautious observation after the resection of primary tumors. Finally, COL1A2 can be reactivated by altering chromatin modifications with methyltransferase and histone deacetylase inhibitors, raising the promise of selective small molecule inhibitors of these enzymes as a potential therapeutic target for HNSCC.

Grant support


Acknowledgments

The authors would like to thank Ms. Yuko Mohri for her excellent technical support.

Conflict of Interest

None declared.

References

K. Misawa et al. / Hypermethylation of COL1A2 in HNSCC


