

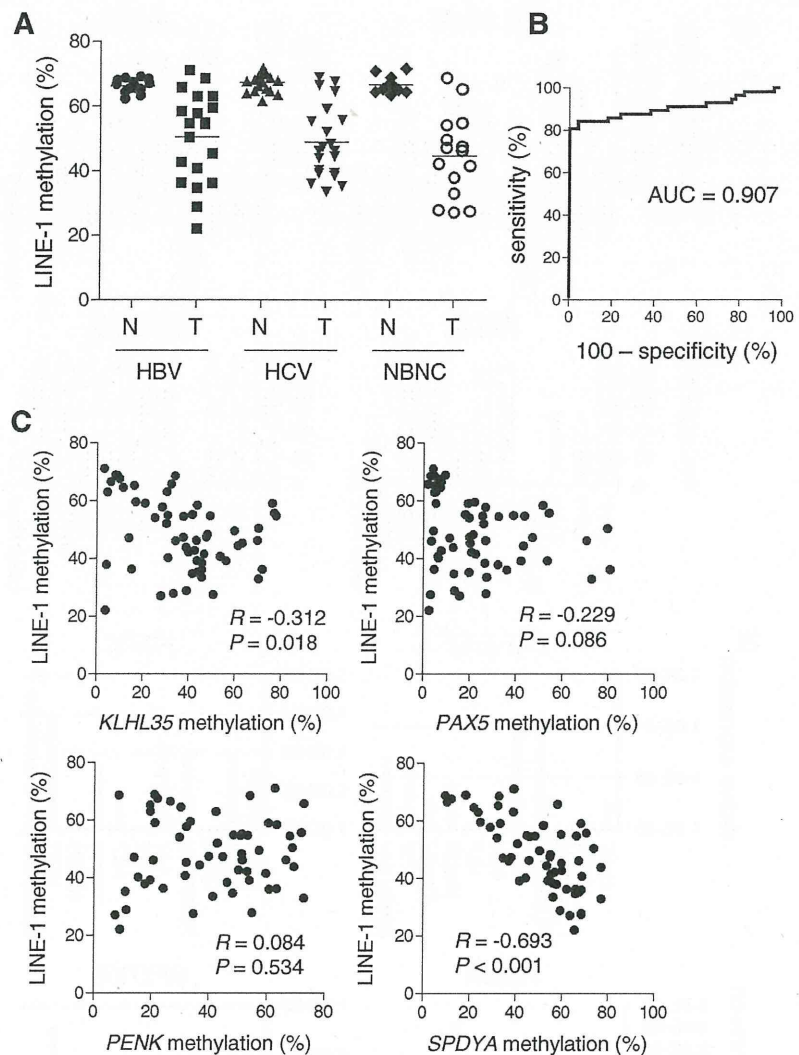
Fig. 3 Analysis of the methylation and expression of the indicated genes in HCC cell lines. **a** Bisulfite pyrosequencing of *KLHL35*, *PAX5*, *PENK*, and *SPDYA* in HCC cell lines and normal liver tissue from a

healthy individual. **b** Quantitative RT-PCR of the four genes in HCC cell lines and normal liver tissue. **c** Quantitative RT-PCR of *KLHL35* and *PAX5* in HCC cell lines, with and without 5-aza-dC (*aza*) treatment

pyrosequencing, we found that levels of LINE-1 methylation were significantly lower in tumor tissues than in their noncancerous counterparts (48.5 vs. 66.8 %, $P < 0.001$). LINE-1 hypomethylation was prevalent, regardless of the tumor's hepatitis virus status, but the average methylation level was lowest in the HBV/HCV-negative tumors (HBV-positive, 50.8 vs. 66.3 %, $P < 0.001$; HCV-positive, 48.9 vs. 67.4 %, $P < 0.001$; HBV/HCV-negative, 44.7 vs. 66.6 %,

$P < 0.001$; Fig. 4a). The ROC curve analysis revealed that LINE-1 methylation discriminated strongly between HCC tissue and noncancerous liver tissue (Fig. 4b), though no significant correlation was found between the levels of LINE-1 methylation and the clinicopathological characteristics of the samples (Table 1). Finally, we tested whether LINE-1 hypomethylation is linked to gene hypermethylation. We found an inverse relationship between the level of

Fig. 4 Analysis of LINE-1 methylation and its association with CpG island hypermethylation in HCC. **a** Summary of bisulfite pyrosequencing analysis of LINE-1 in tumor tissue (*T*) and corresponding noncancerous liver tissue (*N*) from HBV-positive, HCV-positive, and HBV/HCV-negative (*NBNC*) HCC patients. **b** ROC curve analysis of the utility of LINE-1 methylation for distinguishing between HCC tissue and corresponding noncancerous liver tissue from the same HCC patients. **c** Correlation between the level of LINE-1 methylation and methylation of the indicated genes in HCC tissues. The Pearson correlation coefficients and *P* values are shown



LINE-1 methylation and levels of *KLHL35* and *SPDYA* methylation. On the other hand, we found no significant correlation between the LINE-1 hypomethylation and *PAX5* or *PENK* methylation (Fig. 4c).

Discussion

In the present study, we carried out high-throughput CpG island methylation profiling in a set of primary HCC tissues with and without hepatitis virus infection. MCAM analysis enabled us to evaluate the methylation status of more than 6,000 gene promoters with high specificity and sensitivity [13]. Consistent with earlier studies that showed methylation to be more abundant in the HCV-positive HCCs than in the HBV-positive or hepatitis virus-negative HCCs [15, 18], we observed the highest number of methylated genes in HCV-positive HCC tissue. However, we also noted that a

large number of genes were commonly methylated among HCCs, irrespective of the hepatitis virus status, indicating that aberrant methylation of multiple genes may be involved in a common mechanism underlying hepatocarcinogenesis. Moreover, studies have also shown that aberrant methylation detected in tissues or blood samples could be a useful biomarker for early detection of HCC [19, 20]. We therefore validated the methylation status of 14 genes and identified four genes that were frequently methylated in HCC tissues but showed little or no methylation in surrounding noncancerous tissues. The high-tumor specificity suggests that methylation of these genes may not occur at precancerous stages, such as chronic hepatitis or liver cirrhosis; instead, they may be acquired during malignant transformation.

The paired box 5 (*PAX5*) gene is a member of the paired box-containing family of transcription factors, which are involved in the control of organ development and tissue differentiation [21]. *PAX5* is also known to be a B cell-

specific activator protein that plays an essential role during B cell differentiation, neural development, and spermatogenesis. Methylation of the CpG island of *PAX5* was first discovered in breast cancer cells using the MCA technique [22]. Subsequently, methylation and downregulation of *PAX5* were found in lymphoid neoplasms [23]. In addition, while we are preparing the present manuscript, methylation of *PAX5* was reported in HCC and gastric cancer [24, 25]. Restoration of *PAX5* expression in HCC cells induced growth arrest and apoptosis through upregulation of various target genes, including p53, p21, and Fas ligand, suggesting that the *PAX5* acts as a tumor suppressor [24].

The involvement of the kelch-like 35 (*KLHL35*) gene in cancer had not been reported until recently, when a genome-wide analysis of DNA methylation in renal cell carcinoma identified frequent hypermethylation of nine genes, including *KLHL35* [26]. Although the function of the gene product remains unknown, RNAi-induced knockdown of *KLHL35* in HEK293 cells promoted anchorage-independent growth, indicating its possible role in tumorigenesis [26].

The proenkephalin (*PENK*) gene encodes preproenkephalin, a precursor protein that is proteolytically cleaved to produce the endogenous opioid peptides met- and leu-enkephalin. Methylation of the CpG island of *PENK* was first identified in pancreatic cancer cells using the MCA technique [27]. Downregulated expression of *PENK* has also been reported in prostate cancer, suggesting its possible involvement in cancer development [28], and *PENK* methylation was recently identified in lung cancer, bladder cancer, and meningioma [29–31]. Although its functional role in cancer is not fully understood, a recent study showed that in response to cellular stress, *PENK* physically associates with p53 and RelA (p65) and regulates stress-induced apoptosis [32].

The *SPDYA* encodes Spyl, also known as Speedy, an atypical CDK activator known to promote cell survival, prevent apoptosis, and inhibit checkpoint activation in response to DNA damage [33]. The expression of *SPDYA* is upregulated in breast cancer [34], and its overexpression in a mouse model has been shown to accelerate mammary tumorigenesis [35]. Moreover, a recent study showed overexpression of *SYPDA* in HCC and its association with poor prognosis [36]. These results strongly suggest its involvement in oncogenesis. In the present study, we also observed that most of the HCC cell lines tested exhibited greater expression of *SPDYA* than normal liver tissue, regardless of the methylation status. Among the three transcription variants of *SPDYA* annotated in the NCBI Reference Sequence database, transcription start sites of variants 1 and 3 are located within the CpG island, while that of variant 2 are located approximately 5 kb downstream of the CpG island. Thus, the *SPDYA* transcript in HCC cells may be derived from the downstream transcription start site.

By analyzing the LINE-1 methylation levels, we and others have shown that global hypomethylation is a commonly observed feature of HCC [8, 9, 37]. Earlier studies have suggested that the association between global methylation and hepatitis status may be attributable to hepatitis B virus X protein, which can induce aberrant methylation of specific genes and global hypomethylation [38]. By contrast, we found in the present study that LINE-1 hypomethylation is prevalent among HCC tissues, regardless of the hepatitis virus infection, which suggests that global hypomethylation is involved in a common mechanism underlying hepatocarcinogenesis. It has been shown that the timing of global hypomethylation differs among tumor types. For example, hypomethylation is often observed during the early stages of colorectal and gastric carcinogenesis. By contrast, LINE-1 hypomethylation appears to be tumor-specific in HCC; it is rarely found in precancerous lesions such as chronic hepatitis or liver cirrhosis [8, 9]. A recent study showed that global hypomethylation is associated with a poorer prognosis in HCC patients [39]. In addition, the levels of serum LINE-1 hypomethylation in HCC patients reportedly correlate with serum HBs antigen status, large tumor size, and advanced tumor stage [40]. This suggests that hypomethylation may not occur at precancerous stages, and that LINE-1 methylation could be a useful biomarker with which to identify HCC and predict its clinical outcome.

The relationship between LINE-1 hypomethylation and CpG island hypermethylation in cancer is controversial. In one study, LINE-1 methylation levels were reduced in HCCs with the CpG island methylator phenotype, indicating a positive correlation between global hypomethylation and CpG island hypermethylation [9]. Another study showed that LINE-1 hypomethylation was positively correlated with hypermethylation of only a few genes (*p16*, *CACNA1G*, and *CDKN1C*), while methylation of a large number of genes showed inverse or no correlation with LINE-1 hypomethylation [12]. In the present study, we found that methylation of *KLHL35* and *SPYDA* correlates positively with LINE-1 hypomethylation, whereas levels of *PAX5* or *PENK* methylation are independent of LINE-1 methylation. These results suggest that the association between CpG island methylation and global hypomethylation may be site specific, and that hypomethylation of LINE-1 is a more generalized phenomenon than hypermethylation of CpG islands in HCC.

In summary, by screening targets of DNA methylation in HCC, we identified four frequently methylated genes. These genes are methylated in a cancer-specific manner and could be useful molecular markers for diagnosing HCC. In addition, we observed prevalent LINE-1 hypomethylation in HCC, irrespective of hepatitis virus infection. Identification of aberrant methylation in HCC may provide valuable information that not only contributes to our understanding of the pathogenesis

of the disease, but also to the development of new strategies for diagnosis and therapy.

Acknowledgments We thank Dr. Yutaka Kondo for technical advice on MCAM analysis and Masami Ashida for technical assistance. This study was supported in part by a Grant-in-Aid for Scientific Research (B) from the Japan Society for Promotion of Science (Y. Shinomura), a Grant-in-Aid for the Third-term Comprehensive 10-year Strategy for Cancer Control (M. Toyota and H. Suzuki), and a Grant-in-Aid for Cancer Research from the Ministry of Health, Labor, and Welfare, Japan (M. Toyota and H. Suzuki).

Conflicts of interest None

References

- Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005;55:74–108.
- Chen CJ, Yu MW, Liaw YF. Epidemiological characteristics and risk factors of hepatocellular carcinoma. *J Gastroenterol Hepatol.* 1997;12:S294–308.
- Montesano R, Hainaut P, Wild CP. Hepatocellular carcinoma: from gene to public health. *J Natl Cancer Inst.* 1997;89:1844–51.
- Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. *Nat Genet.* 2002;31:339–46.
- Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet.* 2002;3:415–28.
- Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet.* 2009;10:691–703.
- Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res.* 2004;32:e38.
- Takai D, Yagi Y, Habib N, Sugimura T, Ushijima T. Hypomethylation of line1 retrotransposon in human hepatocellular carcinomas, but not in surrounding liver cirrhosis. *Jpn J Clin Oncol.* 2000;30:306–9.
- Kim MJ, White-Cross JA, Shen L, Issa JP, Rashid A. Hypomethylation of long interspersed nuclear element-1 in hepatocellular carcinomas. *Mod Pathol.* 2009;22:442–9.
- Kaneto H, Sasaki S, Yamamoto H, Itoh F, Toyota M, Suzuki H, Ozeki I, Iwata N, Ohmura T, Satoh T, Karino Y, Toyota J, Satoh M, Endo T, Omata M, Imai K. Detection of hypermethylation of the p16(ink4a) gene promoter in chronic hepatitis and cirrhosis associated with hepatitis B or C virus. *Gut.* 2001;48:372–7.
- Takagi H, Sasaki S, Suzuki H, Toyota M, Maruyama R, Nojima M, Yamamoto H, Omata M, Tokino T, Imai K, Shinomura Y. Frequent epigenetic inactivation of sfp genes in hepatocellular carcinoma. *J Gastroenterol.* 2008;43:378–89.
- Lee HS, Kim BH, Cho NY, Yoo EJ, Choi M, Shin SH, Jang JJ, Suh KS, Kim YS, Kang GH. Prognostic implications of and relationship between CpG island hypermethylation and repetitive DNA hypomethylation in hepatocellular carcinoma. *Clin Cancer Res.* 2009;15:812–20.
- Gao W, Kondo Y, Shen L, Shimizu Y, Sano T, Yamao K, Natsume A, Goto Y, Ito M, Murakami H, Osada H, Zhang J, Issa JP, Sekido Y. Variable DNA methylation patterns associated with progression of disease in hepatocellular carcinomas. *Carcinogenesis.* 2008;29:1901–10.
- Arai E, Ushijima S, Gotoh M, Ojima H, Kosuge T, Hosoda F, Shibata T, Kondo T, Yokoi S, Imoto I, Inazawa J, Hirohashi S, Kanai Y. Genome-wide DNA methylation profiles in liver tissue at the precancerous stage and in hepatocellular carcinoma. *Int J Cancer.* 2009;125:2854–62.
- Deng YB, Nagae G, Midorikawa Y, Yagi K, Tsutsumi S, Yamamoto S, Hasegawa K, Kokudo N, Aburatani H, Kaneda A. Identification of genes preferentially methylated in hepatitis C virus-related hepatocellular carcinoma. *Cancer Sci.* 2010;101:1501–10.
- Goto Y, Shinjo K, Kondo Y, Shen L, Toyota M, Suzuki H, Gao W, An B, Fujii M, Murakami H, Osada H, Taniguchi T, Usami N, Kondo M, Hasegawa Y, Shimokata K, Matsuo K, Hida T, Fujimoto N, Kishimoto T, Issa JP, Sekido Y. Epigenetic profiles distinguish malignant pleural mesothelioma from lung adenocarcinoma. *Cancer Res.* 2009;69:9073–82.
- Suzuki H, Yamamoto E, Nojima M, Kai M, Yamano HO, Yoshikawa K, Kimura T, Kudo T, Harada E, Sugai T, Takamaru H, Niinuma T, Maruyama R, Yamamoto H, Tokino T, Imai K, Toyota M, Shinomura Y. Methylation-associated silencing of microRNA-34b/c in gastric cancer and its involvement in an epigenetic field defect. *Carcinogenesis.* 2010;31:2066–73.
- Nishida N, Nagasaka T, Nishimura T, Ikai I, Boland CR, Goel A. Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. *Hepatology.* 2008;47:908–18.
- Wong IH, Zhang J, Lai PB, Lau WY, Lo YM. Quantitative analysis of tumor-derived methylated p16ink4a sequences in plasma, serum, and blood cells of hepatocellular carcinoma patients. *Clin Cancer Res.* 2003;9:1047–52.
- Zhang YJ, Wu HC, Shen J, Ahsan H, Tsai WY, Yang HI, Wang LY, Chen SY, Chen CJ, Santella RM. Predicting hepatocellular carcinoma by detection of aberrant promoter methylation in serum DNA. *Clin Cancer Res.* 2007;13:2378–84.
- Carotta S, Holmes ML, Pridans C, Nutt SL. Pax5 maintains cellular identity by repressing gene expression throughout B cell differentiation. *Cell Cycle.* 2006;5:2452–6.
- Palmisano WA, Crume KP, Grimes MJ, Winters SA, Toyota M, Esteller M, Joste N, Baylin SB, Belinsky SA. Aberrant promoter methylation of the transcription factor genes pax5 alpha and beta in human cancers. *Cancer Res.* 2003;63:4620–5.
- Lazzi S, Bellan C, Onnis A, De Falco G, Sayed S, Kostopoulos I, Onorati M, D'Amuri A, Santopietro R, Vindigni C, Fabbri A, Righi S, Pileri S, Tosi P, Leoncini L. Rare lymphoid neoplasms coexpressing B- and T-cell antigens. The role of pax-5 gene methylation in their pathogenesis. *Hum Pathol.* 2009;40:1252–61.
- Liu W, Li X, Chu ES, Go MY, Xu L, Zhao G, Li L, Dai N, Si J, Tao Q, Sung JJ, Yu J. Paired box gene 5 is a novel tumor suppressor in hepatocellular carcinoma through interaction with p53 signaling pathway. *Hepatology.* 2011;53:843–53.
- Li X, Cheung KF, Ma X, Tian L, Zhao J, Go MY, Shen B, Cheng AS, Ying J, Tao Q, Sung JJ, Kung HF, Yu J. Epigenetic inactivation of paired box gene 5, a novel tumor suppressor gene, through direct upregulation of p53 is associated with prognosis in gastric cancer patients. *Oncogene.* 2011. doi:10.1038/onc.2011.511.
- Morris MR, Ricketts CJ, Gentle D, McRonald F, Carli N, Khalili H, Brown M, Kishida T, Yao M, Banks RE, Clarke N, Latif F, Maher ER. Genome-wide methylation analysis identifies epigenetically inactivated candidate tumour suppressor genes in renal cell carcinoma. *Oncogene.* 2011;30:1390–401.
- Ueki T, Toyota M, Skinner H, Walter KM, Yeo CJ, Issa JP, Hruban RH, Goggins M. Identification and characterization of differentially methylated CpG islands in pancreatic carcinoma. *Cancer Res.* 2001;61:8540–6.
- Goo YA, Goodlett DR, Pascal LE, Worthington KD, Vessella RL, True LD, Liu AY. Stromal mesenchyme cell genes of the human prostate and bladder. *BMC Urol.* 2005;5:17.

29. Chung JH, Lee HJ, Kim BH, Cho NY, Kang GH. DNA methylation profile during multistage progression of pulmonary adenocarcinomas. *Virchows Arch.* 2011;459:201–11.
30. Chung W, Bondaruk J, Jelinek J, Lotan Y, Liang S, Czerniak B, Issa JP. Detection of bladder cancer using novel DNA methylation biomarkers in urine sediments. *Cancer Epidemiol Biomarkers Prev.* 2011;20:1483–91.
31. Kishida Y, Natsume A, Kondo Y, Takeuchi I, An B, Okamoto Y, Shinjo K, Saito K, Ando H, Ohka F, Sekido Y, Wakabayashi T. Epigenetic subclassification of meningiomas based on genome-wide DNA methylation analyses. *Carcinogenesis.* 2012;33:436–41.
32. McTavish N, Copeland LA, Saville MK, Perkins ND, Spruce BA. Proenkephalin assists stress-activated apoptosis through transcriptional repression of nf-kappab- and p53-regulated gene targets. *Cell Death Differ.* 2007;14:1700–10.
33. Gastwirt RF, McAndrew CW, Donoghue DJ. Speedy/ringo regulation of CDKs in cell cycle, checkpoint activation and apoptosis. *Cell Cycle.* 2007;6:1188–93.
34. Zucchi I, Mento E, Kuznetsov VA, Scotti M, Valsecchi V, Simionati B, Vicinanza E, Valle G, Pilotti S, Reinbold R, Vezzoni P, Albertini A, Dulbecco R. Gene expression profiles of epithelial cells microscopically isolated from a breast-invasive ductal carcinoma and a nodal metastasis. *Proc Natl Acad Sci U S A.* 2004;101:18147–52.
35. Golipour A, Myers D, Seagraves T, Murphy D, Evan GI, Donoghue DJ, Moorehead RA, Porter LA. The *spy1/ringo* family represents a novel mechanism regulating mammary growth and tumorigenesis. *Cancer Res.* 2008;68:3591–600.
36. Ke Q, Ji J, Cheng C, Zhang Y, Lu M, Wang Y, Zhang L, Li P, Cui X, Chen L, He S, Shen A. Expression and prognostic role of *spy1* as a novel cell cycle protein in hepatocellular carcinoma. *Exp Mol Pathol.* 2009;87:167–72.
37. Lin CH, Hsieh SY, Sheen IS, Lee WC, Chen TC, Shyu WC, Liaw YF. Genome-wide hypomethylation in hepatocellular carcinogenesis. *Cancer Res.* 2001;61:4238–43.
38. Park IY, Sohn BH, Yu E, Suh DJ, Chung YH, Lee JH, Surzycki SJ, Lee YI. Aberrant epigenetic modifications in hepatocarcinogenesis induced by hepatitis B virus X protein. *Gastroenterology.* 2007;132:1476–94.
39. Calvisi DF, Simile MM, Ladu S, Pellegrino R, De Murtas V, Pinna F, Tomasi ML, Frau M, Viridis P, De Miglio MR, Muroli MR, Pascale RM, Feo F. Altered methionine metabolism and global DNA methylation in liver cancer: relationship with genomic instability and prognosis. *Int J Cancer.* 2007;121:2410–20.
40. Tangkijvanich P, Hourpai N, Rattanananyong P, Wisedopas N, Mahachai V, Mutirangura A. Serum *line-1* hypomethylation as a potential prognostic marker for hepatocellular carcinoma. *Clin Chim Acta.* 2007;379:127–33.

Long noncoding RNA involvement in cancer

Reo Maruyama & Hiromu Suzuki*

Department of Molecular Biology, Sapporo Medical University, Sapporo 060-8556, Japan

Recent advances in genome and transcriptome analysis have enabled identification of numerous members of a new class of noncoding RNA, long noncoding RNA (lncRNA). lncRNAs are broadly defined as RNA molecules greater than 200 nt in length and lacking an open reading frame. Recent studies provide evidence that lncRNAs play central roles in a wide range of cellular processes through interaction with key component proteins in the gene regulatory system, and that alteration of their cell- or tissue-specific expression and/or their primary or secondary structures is thought to promote cell proliferation, invasion and metastasis. The biological and molecular characteristics of the large majority of lncRNAs remains unknown, and it is anticipated that improved understanding of the roles played by lncRNAs in cancer will lead to the development of novel biomarkers and effective therapeutic strategies. [BMB Reports 2012; 45(11): 604-611]

INTRODUCTION

Recent advances in genome analysis, including microarray and massively parallel sequencing, have shown that a much larger portion of the human genome is pervasively transcribed into RNA than previously recognized. Moreover, much of the evidence emerging in recent years has highlighted the biological and pathological importance of RNA molecules that lack protein-coding potential; these are collectively referred to as noncoding RNAs (ncRNAs) (1). Long ncRNAs (lncRNAs) are broadly defined as transcribed RNA molecules greater than 200 nt in length and lacking an open reading frame of significant length (less than 100 amino acids). Although there are no specific definitions, lncRNAs can be categorized into several subgroups based on their locations and characteristics. For instance, antisense RNAs are transcribed from the opposite strand of a protein-coding gene, while lncRNAs transcribed from intergenic regions are referred to as large intergenic

ncRNAs (lincRNAs).

Much about the molecular and biological characteristics of lncRNAs is as yet unknown, but what is known suggests that many show expression patterns that are spatially and temporally specific and are generally poorly conserved among species. Cabili *et al.* recently used integrated RNA-seq data to construct a reference catalog of 8195 human lincRNAs (2). According to their reports, lincRNAs, like protein-coding transcripts, are transcribed by RNA polymerase II and are spliced and polyadenylated, but the maximum expression levels of lincRNAs are 10 times lower than those of protein-coding transcripts. lincRNAs are also smaller in size than protein-coding RNAs and have fewer exons: on average, lincRNAs are ~1 kb in length with 2.9 exons, whereas protein-coding transcripts are ~2.9 kb in length with 10.7 exons. In addition, lincRNAs are alternatively spliced more frequently than protein-coding mRNAs (2.3 isoforms per lincRNA locus, on average). It is also noteworthy that whereas 78% of examined lincRNAs exhibit tissue-specific expression patterns, only 19% of protein-coding transcripts do so. In fact, the tissue (cell or context) specificity of lincRNAs expression has been reported often enough (3, 4) to suggest that it is an important feature of lincRNAs.

Although the specific functions of the large majority of lncRNAs remain unknown, recent studies have begun to shed light on the critical roles played by these molecules in a variety of cellular processes, including differentiation, development and tumorigenesis. In this review, we will briefly outline the known functions of lncRNAs and their involvement in cancer.

FUNCTIONS OF lncRNA

Although, as mentioned, the function of most lncRNAs remains unknown, dozens of examples of biologically functional lncRNAs have already been reported, and the number of such examples is rapidly increasing (Table 1). Basically, lncRNAs appear to be involved in all aspects of gene regulation, including chromosome dosage-compensation, imprinting, epigenetic regulation, nuclear and cytoplasmic trafficking, transcription, mRNA splicing and translation. Through such gene regulation, lncRNAs are involved in a wide range of biological processes, including proliferation, cell cycle, apoptosis, differentiation and maintenance of pluripotency, among others. There is a good review of lncRNA's functions, focusing on biological and pathological processes in cancer (5); refer to that article for further details. On the other hand, it is also useful to look at

*Corresponding author. Tel: +81-11-611-2111; Fax: +81-11-622-1918; E-mail: hsuzuki@sapmed.ac.jp
<http://dx.doi.org/10.5483/BMBRep.2012.45.11.227>

Received 10 October 2012

Keywords: Biomarker, Cancer, Gene regulation, Histone modification, LincRNA, Long noncoding RNA, Oncogene, Tumor suppressor

Table 1. Representative lncRNAs implicated in cancer

Name	Size	Cancer type	Biological function	Molecular function	References
Oncogenic					
HOTAIR	2.2 kb	Breast, hepatocellular, colorectal, pancreatic, GIST	Promotes invasion and metastasis, Modulates cancer epigenome	Scaffold (PRC2, LSD1), Guide (<i>in trans</i>)	(11, 26-29)
ANRIL	2.2 kb	Prostate, leukemia, neural system, melanoma	Suppresses senescence via INK4A	Scaffold (PRC1, PRC2), Guide (<i>in cis</i>)	(30-35)
MALAT1	8.7 kb	Lung, prostate, breast, colon, hepatocellular	Regulates alternative splicing of pre-mRNA	Scaffold (nuclear paraspeckle)	(36-40, 47)
PCAT-1	1.9 kb	Prostate	Promotes cell proliferation, inhibits BRCA2	Unknown	(42)
PCGEM1	1.6 kb	Prostate	Inhibits apoptosis, promotes cell proliferation	Unknown	(48, 49)
TUC338	0.6 kb	Hepatocellular	Promotes cell proliferation	Unknown	(50)
uc.73a	0.2 kb	Leukemia, colorectal	Inhibits apoptosis, promotes cell proliferation	Unknown	(51)
SPRY4-IT1	0.7 kb	Melanoma	Promotes cell proliferation and invasion, Inhibits apoptosis	Unknown	(52)
ncRAN	2.3 kb	Neuroblastoma, bladder	Promotes cell proliferation and invasion	Unknown	(53, 54)
PRNCR1	13 kb	Prostate	Promotes cell proliferation	Unknown	(55)
H19	2.3 kb	Breast, hepatocellular	Promotes cell proliferation, both oncogenic and tumor suppressive functions reported	Unknown	(56, 57)
Tumor suppressive					
GAS5	0.6-1.8 kb	Breast	Induces growth arrest and apoptosis	Decoy (glucocorticoid receptor)	(21, 44)
MEG3	1.6-1.8 kb	Meningioma, hepatocellular, leukemia, pituitary	Mediates p53 signaling	Unknown	(58, 59)
PTENP1	3.9 kb	Prostate, colon	Inhibits cell proliferation	Decoy (PTEN-suppressing miRNAs)	(23)
lincRNA-p21	3.1 kb	Mouse models of lung, sarcoma, lymphoma	Induces apoptosis by repressing p53 targets	Guide (hnRNP-k, <i>in trans</i>)	(10)
Unknown, Biomarker					
HULC	0.5 kb	Hepatocellular	Unknown	Decoy (miR-372)	(60, 61)
PCA3	3.6 kb	Prostate	Unknown	Unknown	(62)
PANDA	1.5 kb	Unknown	Inhibits apoptosis by sequestering NF-YA	Decoy (NF-YA)	(4)

lncRNAs in the context of their molecular functionality (6). In this review we will focus on three molecular functions of lncRNAs: guide, scaffold and decoy.

Guide

One group of lncRNAs is able to bind specific proteins and then direct the localization of the resultant complex to specific targets. Such lncRNAs can guide proteins either *in cis* (on neighboring genes) or *in trans* (on distantly located genes). XIST, one of the most well-studied lncRNAs, guides PRC2 (polycomb repressive complex 2) to one of the two X chromosomes *in cis* to achieve X inactivation (7). Two other examples of lncRNAs that function as guides *in cis* are AIR and HOTTIP. AIR silences transcription of its target gene on the paternal chromosome by recruiting G9a and then mediating targeted histone H3 lysine 9 (H3K9) methylation and allelic silencing (8). HOTTIP, which is transcribed from the end of HOXA cluster, binds to WDR5 and recruits the MLL histone H3 lysine 4 (H3K4) methyltransferase complex to the HOXA cluster to sup-

port active chromatin confirmation (9).

Some lncRNAs, including lincRNA-p21 and HOTAIR, are able to alter and regulate epigenetic states and gene expression across multiple sites *in trans* (10, 11). lincRNA-p21 is transcribed upstream of CDKN1A gene and acts as a transcriptional repressor through its interaction with hnRNP-K, which it guides to target sites. Knocking down lincRNA-p21 alters the expression of over 1,000 genes, suggesting lncRNAs may be able to regulate numerous genes *in trans*. In addition, some lncRNAs may serve as cellular "navigation systems" for proteins lacking direct DNA binding capacity (12).

The mechanism by which lncRNAs specifically regulate their target genes remains unclear, and their binding sites throughout the genome are largely unknown. In one recent study, Chu and colleagues addressed this question using a novel assay they named ChIRP (Chromatin Isolation by RNA Purification)-seq, which is a method for genome-wide mapping of lncRNA binding sites *in vivo* (13). This analysis enabled them to obtain a high-resolution map of ncRNA occupancy

throughout the genome and to identify a set of 832 HOTAIR binding sites in human breast cancer cells. Interestingly, binding sites for HOTAIR are focal (<500 bp) and located in the midst of a broad polycomb binding domain, which suggests HOTAIR may act as a pioneering factor able to recruit polycomb to its target genes and then bilaterally spread the repressive regions outward. They also discovered an underlying DNA sequence motif enriched in HOTAIR binding sites, indicating the existence of a new class of regulatory element: lncRNA target sites. We now think that lncRNAs can function as sequence-specific transcription factors.

Scaffold

Another class of lncRNAs may possess distinct domains that bind different effector molecules. Such lncRNAs can mediate assembly of multiple molecular components in temporally and spatially specific manner. Two examples of lncRNAs that act as scaffolds are ANRIL, which interacts with components from PRC1 (polycomb repressive complex 1) and PRC2 (14, 15), and KCNQ1OT1, which binds both G9a and PRC2 (16). In addition, MALAT1 and NEAT1 serve as molecular scaffolds for proteins within nuclear speckles and paraspeckles, respectively (17). Depletion of NEAT1 is sufficient to cause loss of paraspeckles from within the nucleus, and overexpressing NEAT1, but not paraspeckle-associated proteins, leads to an increase in the number of paraspeckles, suggesting NEAT1 plays an essential role as a scaffold in the formation of paraspeckles.

A recent study revealed that HOTAIR interacts with two chromatin modifying complexes, the PRC2 complex ("writer" of a repressive mark, H3K27 trimethylation) and the LSD1/CoREST H3K4 demethylase complex ("eraser" of an activating mark, H3K4 trimethylation) (18). Using a series of deletion mutants, the PRC2 binding domain was mapped to the 5' end (the first 300 nt) of HOTAIR, while the LSD1 binding site corresponds to the 3' end. This suggests that HOTAIR that bridges between the PRC2 and LSD1 complexes, and that the resultant HOTAIR/PRC2/LSD1 complex can suppress gene expression via multiple mechanisms.

This finding is not applicable only to HOTAIR; many other lncRNAs also appear to interact with both the PRC2 and LSD1 complexes. For example, Khalil and colleagues performed RIP-chip assays (RNA coimmunoprecipitation combined with high throughput lincRNA microarray) using antibodies directed against several proteins involved in chromatin modifying complexes (PRC2, CoREST and SMCX) (19). They found that as many as 38% of lincRNAs expressed in the cell types studied reproducibly associate with one of these complexes. In mouse embryonic stem cells, moreover, a number of lincRNAs were found to be strongly associated with multiple chromatin modifier complexes (20). For example, eight lincRNAs bind to the PRC2 H3K27 and ESET H3K9 methyltransferase complexes and the JARID1C H3K4 demethylase complex. Similarly, 17 lincRNAs were found to bind to the PRC2, PRC1 and JARID1B complexes. Taken together, these results suggest the attractive

hypothesis that lincRNAs bind to ubiquitously expressed chromatin modifying complexes in order to guide them to specific genomic regions.

Decoy

A third class of lncRNAs bind and then sequester a protein or RNA target, but do not exert additional effects. By acting as molecular decoys, these lncRNAs negatively regulate the expression of their targets. Examples of lncRNAs with decoy functionality include GAS5 (growth arrest-specific 5), which binds to the glucocorticoid receptor (GR) and represses GR-induced genes (21); PANDA (P21 associated ncRNA DNA damage activated), which binds to the transcription factor NF-YA to negatively regulate expression of pro-apoptotic genes (4); and TERRA (telomeric repeat-containing RNA), which interacts with telomerase via a repeat sequence to reduce enzyme activity (22).

Interestingly, several pseudogenes also reportedly act as molecular decoys. The 3'UTR of PTENP1, a tumor suppressor pseudogene, was found to bind the same set of regulatory miRNA sequences that normally target the tumor-suppressor gene PTEN, which reduces the downregulation of PTEN mRNA and allows its translation into the tumor-suppressor protein PTEN (23). It was also shown that the PTENP1 locus is selectively lost in human cancer, and that similar relationships also exist between other cancer-related genes and their pseudogenes (24). These findings attribute a novel biological function to expressed pseudogenes, as they can regulate coding gene expression and reveal a non-coding function in mRNAs.

DYSREGULATION OF lncRNAs IN HUMAN CANCER

HOTAIR

The lincRNA HOTAIR was originally discovered by Rinn and colleagues (25). It is encoded within the HOXC gene cluster and acts *in trans* to regulate HOXD genes through recruitment of PRC2 to induce trimethylation of H3K27 (H3K27me3). Remarkably, pull-down assays with PRC2 components demonstrated a direct and specific interaction with HOTAIR. This observation that HOTAIR binds PRC2 and induces epigenetic silencing of another HOX cluster on a different chromosome was an unexpected and novel finding.

In breast cancer, elevated expression of HOTAIR reportedly correlates with a poor prognosis and tumor metastasis (11). It is noteworthy that expression of a single lincRNA in primary tumors can be a powerful predictor of eventual metastasis and death. Enforced expression of HOTAIR induces genome-wide re-targeting of PRC2, leading to altered H3K27me3 and gene expression, and increased cancer invasiveness and metastasis. The link between HOTAIR and metastatic disease depends on both the direct interaction between the ncRNA and its protein partner and between the ncRNA and its target DNA sequence. Furthermore, several other studies have also shown that the level of HOTAIR expression correlates positively with metastasis and poor outcome in hepatocellular carcinoma, color-

ectal cancer and pancreatic cancer (26-28). We also discovered that upregulation of HOTAIR is strongly associated with aggressiveness in gastrointestinal stromal tumors (GISTs) (29). Interestingly, in malignant GISTs, HOTAIR is concurrently overexpressed with collinear HOXC genes and an oncogenic microRNA, miR-196a. We also observed enrichment of an active histone mark, H3K4me3, over a wide range of the HOXC cluster, suggesting the entire region is epigenetically activated in malignant GISTs. Taken together, these results suggest that lincRNAs play active roles in modulating the cancer epigenome and could be useful targets for cancer diagnosis and therapy.

ANRIL

ANRIL (antisense lincRNA of the INK4 locus) is transcribed antisense to The INK4 locus and mediates INK4a transcriptional repression *in cis* (30). Independent studies have shown that overexpression of ANRIL in prostate cancer results in the silencing of INK4b/ARF/INK4a and p15/CDKN2B due to heterochromatin formation (14, 31). ANRIL interacts with SUZ12 (suppressor of zeste 12 homolog), a subunit of PRC2 (15) and with CBX7 (chromo-box homolog 7), a subunit of PRC1 (14). ANRIL may recruit multiple sets of chromatin-modifying complexes to a target gene for silencing, serving as a molecular scaffold. Genome-wide association studies (GWAS) have linked ANRIL with increased susceptibility to coronary disease, intracranial aneurysm and type 2 diabetes, as well as to several types of cancer, including acute lymphoblastic leukemia, glioma, basal cell carcinoma, nasopharyngeal carcinoma, breast cancer and plexiform neurofibromas (32-35).

MALAT1

MALAT1 (Metastasis-Associated Lung Adenocarcinoma Transcript 1) was initially identified in a screen for genes associated with metastasis (36). It is abundantly expressed in many human cell types and is highly conserved across several species. MALAT1 is reportedly upregulated in multiple malignancies, including lung cancer, uterine endometrial stromal sarcoma, cervical cancer and hepatocellular carcinoma (36-38). In lung metastasizing tumors, MALAT1 expression is three-fold higher than in non-metastasizing tumors (36), and it can serve as an independent prognostic parameter for patient survival in early stage lung adenocarcinoma (39). MALAT1 promotes the motility of lung cancer cells through transcriptional or post-transcriptional regulation of motility-related genes (40).

MALAT1 localizes to nuclear speckles, which contain several proteins known to be involved in alternative splicing (41). It appears that MALAT1 forms a molecular scaffold for several of the proteins present within nuclear speckles, and modulates the phosphorylation of SR proteins. Depletion of MALAT1 is sufficient to alter the patterns of alternative splicing of a subset of mRNAs.

LincRNA-p21

Recent studies identified a number of lincRNAs induced by the p53 tumor-suppressor gene (10). LincRNA-p21 is located upstream of the CDKN1A gene on mouse chromosome 17 and is directly activated by p53 in response to DNA damage. LincRNA-p21 acts as a transcriptional repressor in the canonical p53 pathway and plays a role in triggering apoptosis. Inhibition of lincRNA-p21 alters the expression of hundreds of genes normally repressed by p53, potentially explaining how p53 can activate large numbers of genes while simultaneously repressing many others. LincRNA-p21 also interacts with heterogeneous nuclear ribonucleoprotein K (hnRNP-K), a well-known RNA binding protein that acts as a transcriptional repressor. A 780-nt region at the 5' end of lincRNA-p21 is necessary for interaction with hnRNP-K, which is required for proper genomic localization of hnRNP-K at repressed genes and for regulation of p53-mediated apoptosis. Although lincRNA-p21 has not been directly associated with disease, we would speculate that loss of lincRNA-p21 function could be an important factor contributing to cancer initiation.

PANDA

PANDA is also induced in a p53-dependent manner. After DNA damage, p53 directly binds to the CDKN1A locus to activate PANDA (4), which appears to possess decoy function. PANDA inhibits the expression of apoptotic genes by directly binding to and sequestering the transcription factor NF-YA away from target gene promoters. PANDA is overexpressed in a subset of human breast cancers, and its depletion can sensitize cells to chemotherapeutic agents.

PCAT-1

Prensner and colleagues used high throughput RNA-Seq with a large panel of clinical samples to comprehensively evaluate the ncRNAs dysregulated in prostate cancer (42). They identified approximately 1,800 lincRNAs in prostate tissue, of which 121 were transcriptionally dysregulated in prostate cancer. Among them, PCAT-1 (prostate cancer associated transcript -1) showed tissue-specific expression and was selectively upregulated in prostate cancer. Like HOTAIR, PCAT-1 functions predominantly as a transcriptional repressor by facilitating trans-regulation of genes preferentially involved in mitosis and cell division, including known tumor suppressor genes such as BRCA2. The discovery of PCAT-1 highlights the usefulness of unbiased transcriptome analysis when investigating the actions of lincRNAs in cancer.

GAS5

GAS5 was originally identified as a gene highly expressed in cells whose growth was arrested (43). The human GAS5 gene is transcribed from chromosome 1q25.1 and is alternatively spliced. GAS5 sensitizes cells to apoptosis by regulating the activity of glucocorticoids in response to nutrient starvation (21). GAS5 binds to the DNA-binding domain of the GR,

where it acts as a decoy preventing the GR from interacting with cognate glucocorticoid response elements (GRE). The binding of GAS5 to the GR is sufficient to repress GR-induced genes such as *clAP2*.

Implicating GAS5 in breast cancer is the observation that levels of GAS5 transcript are significantly lower in breast cancer cells than in unaffected normal breast epithelium (44). In addition, genetic aberrations at the GAS5 locus have been found in many types of tumors, including melanomas and breast and prostate cancers (45, 46), though their functional significance has not yet been established. The GAS5 gene locus has also been linked to increased susceptibility to autoimmune disorders such as systemic lupus erythematosus in the mouse BXSB strain (21). Chromosomal translocations affecting the 1q25 locus containing the GAS5 gene have been detected in melanoma, B-cell lymphoma, and prostate and breast cancer.

CONCLUDING REMARKS

In recent years, technological advances in high-throughput sequencing have enabled us to identify numerous ncRNAs, which has improved our appreciation for the complexity of transcriptome and enabled us to demonstrate the dysregulated expression of a number of lncRNAs in various types of cancer. The affected molecules are now thought to function as oncogenes or tumor suppressors. However, the biological and molecular characteristics of most lncRNAs remain unknown, and much effort will be needed before a full understanding of their roles in normal and cancer cells is attained. This work will in part entail bioinformatics analysis, and secondary structure prediction will be important for identifying potentially functional motifs in lncRNAs. Elucidation of the structures of genes encoding lncRNAs, including the promoters, transcription start sites and enhancers will be essential for understanding the mechanisms governing their spatiotemporal expression patterns. Moreover, greater knowledge of the gene structures will lead to complete cloning of lncRNAs, which will facilitate functional studies using expression constructs and mouse models. Analysis of genetic and epigenetic alterations of lncRNA genes will also provide clues to understanding their pathological roles, and the mechanisms by which such alterations affect lncRNA function must be carefully characterized. Although the roles played by lncRNAs in cancer have just begun to be revealed, it is anticipated that advances in the study of lncRNAs will yield new diagnostic biomarkers as well as RNA-based therapeutic strategies.

Acknowledgements

We thank Dr. William Goldman for editing the manuscript. H. Suzuki is supported by a Grant-in-Aid for Scientific Research (C) from the Japan Society for Promotion of Science, a Grant-in-Aid for the Third-term Comprehensive 10-year Strategy for Cancer Control from the Ministry of Health, Labor, and Welfare, Japan and the A3 Foresight Program from the

Japan Society for Promotion of Science.

REFERENCES

1. Esteller, M. (2011) Non-coding RNAs in human disease. *Nat. Rev. Genet.* **12**, 861-874.
2. Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A. and Rinn, J. L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915-1927.
3. Mercer, T. R., Dinger, M. E., Sunken, S. M., Mehler, M. F. and Mattick, J. S. (2008) Specific expression of long non-coding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 716-721.
4. Hung, T., Wang, Y., Lin, M. F., Koegel, A. K., Kotake, Y., Grant, G. D., Horlings, H. M., Shah, N., Umbricht, C., Wang, P., Kong, B., Langerod, A., Borresen-Dale, A. L., Kim, S. K., van de Vijver, M., Sukumar, S., Whitfield, M. L., Kellis, M., Xiong, Y., Wong, D. J. and Chang, H. Y. (2011) Extensive and coordinated transcription of non-coding RNAs within cell-cycle promoters. *Nat. Genet.* **43**, 621-629.
5. Gutschner, T. and Diederichs, S. (2012) The Hallmarks of Cancer: A long non-coding RNA point of view. *RNA Biol.* **9**, 703-719. doi: 10.4161/rna.20481.
6. Wang, K. C. and Chang, H. Y. (2011) Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43**, 904-914.
7. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. and Lee, J. T. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750-756.
8. Nagano, T., Mitchell, J. A., Sanz, L. A., Pauler, F. M., Ferguson-Smith, A. C., Feil, R. and Fraser, P. (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**, 1717-1720.
9. Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B. R., Protacio, A., Flynn, R. A., Gupta, R. A., Wysocka, J., Lei, M., Dekker, J., Helms, J. A. and Chang, H. Y. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120-124.
10. Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M. J., Kenzelmann-Broz, D., Khalil, A. M., Zuk, O., Amit, I., Rabani, M., Attardi, L. D., Regev, A., Lander, E. S., Jacks, T. and Rinn, J. L. (2010) A large intergenic non-coding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**, 409-419.
11. Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S. and Chang, H. Y. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-1076.
12. Niland, C. N., Merry, C. R. and Khalil, A. M. (2012) Emerging roles for long non-coding RNAs in cancer and neurological disorders. *Front Genet.* **3**, 25.
13. Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. and Chang, H. Y. (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions.