Cancer Letters 322 (2012) 204–212

Contents lists available at SciVerse ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Development of a novel approach, the epigenome-based outlier approach, to identify tumor-suppressor genes silenced by aberrant DNA methylation

Mizuho Kikuyama [a,b], Hideyuki Takeshima [a], Takayuki Kinoshita [c], Eriko Okochi-Takada [a], Mika Wakabayashi [a], Sadako Akashi-Tanaka [c], Toshihisa Ogawa [b], Yasuyuki Seto [b], Toshikazu Ushijima [a,*]

[a] Division of Epigenomics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, 104-0045 Tokyo, Japan
[b] Department of Metabolic Care and Endocrine Surgery, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 113-8655 Tokyo, Japan
[c] Breast Surgery Division, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, 104-0045 Tokyo, Japan

ARTICLE INFO

Article history:
Received 4 January 2012
Received in revised form 10 March 2012
Accepted 12 March 2012

Keywords:
Epigenetics
Aberrant DNA methylation
Tumor-suppressor gene

ABSTRACT

Identification of tumor-suppressor genes (TSGs) silenced by aberrant methylation of promoter CpG islands (CGIs) is important, but hampered by a large number of genes methylated as passengers of carcinogenesis. To overcome this issue, we here took advantage of the fact that the vast majority of genes methylated in cancers lack, in normal cells, RNA polymerase II (Pol II) and have trimethylation of histone H3 lysine 27 (H3K27me3) in their promoter CGIs. First, we demonstrated that three of six known TSGs in breast cancer and two of three in colon cancer had Pol II and lacked H3K27me3 in normal cells, being outliers to the general rule. *BRCA1*, *HOXA5*, *MLH1*, and *RASSF1A* had high Pol II, but were expressed only at low levels in normal cells, and were unlikely to be identified as outliers by their expression statuses in normal cells. Then, using epigenome statuses (Pol II binding and H3K27me3) in normal cells, we made a genome-wide search for outliers in breast cancers, and identified 14 outlier promoter CGIs. Among these, *DZIP1*, *FBN2*, *HOXA5*, and *HOXC9* were confirmed to be methylated in primary breast cancer samples. Knockdown of *DZIP1* in breast cancer cell lines led to increases of their growth, suggesting it to be a novel TSG. The outliers based on their epigenome statuses contained unique TSGs, including *DZIP1*, compared with those identified by the expression microarray data. These results showed that the epigenome-based outlier approach is capable of identifying a different set of TSGs, compared to the expression-based outlier approach.

1. Introduction

Aberrant DNA methylation of promoter CpG islands (CGIs), especially that of nucleosome-free regions (NFRs), causes silencing of their downstream genes [1], and is known as an alternative mechanism to point mutations and chromosomal deletion for inactivation of tumor-suppressor genes (TSGs) [2–6]. Now, TSGs involved in various biological pathways such as cell proliferation, cell adhesion, and DNA repair are known to be silenced by aberrant DNA methylation in human cancers [7,8]. To identify novel TSGs silenced by aberrant DNA methylation, two major approaches for genome-wide screening have been reported, namely (1) genome-wide screening of methylated genes and (2) screening of genes whose expression is induced by a DNA demethylating agent (chemical genomic screening) [9–13].

A genome-wide screening of methylated genes usually isolates several hundreds to one thousand genes whose promoter CGIs are methylated in cancers [12,14]. Unfortunately, most of these genes are not expressed or expressed at only low levels in normal cells [14]. Since genes with low transcription tend to be methylated [15], most of the genes identified by genome-wide screening of methylated genes are considered to be methylated as a consequence of carcinogenesis. In chemical genomic screening, cell lines are treated with a DNA demethylating agent, 5-aza-2′-deoxycytidine (5-aza-dC), and genes whose expression is induced are screened by expression microarray analysis. However, since 5-aza-dC forms DNA adducts [16] and induces double-strand breaks [17], treatment with 5-aza-dC leads to up-regulation of not only methylation-silenced genes but also numerous genes involved in responses to DNA damage, including p53-activated genes [18–21].

It was recently shown that the vast majority of genes aberrantly methylated in cancers have specific epigenetic statuses in normal cells, namely the lack of RNA polymerase II (Pol II) and the presence of trimethylation of histone H3 lysine 27 (H3K27me3) [14,22]. H3K27me3 is known to be most influential on the susceptibility of genes to become methylated among various histone

Abbreviations: CGIs, CpG islands; ChIP, chromatin immunoprecipitation; FHC, normal fetal human colon epithelial cell line; HMECs, normal human mammary epithelial cells; H3K27me3, trimethylation of histone H3 lysine 27; NFR, nucleosome-free region; Pol II, RNA polymerase II; TSG, tumor-suppressor gene.

* Corresponding author. Fax: +81 3 5565 1753.
E-mail address: tushijim@ncc.go.jp (T. Ushijima).

0304-3835/$ - see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.canlet.2012.03.016

modifications [14]. At the same time, approximately 5% of the genes methylated in cancers do not follow the general rule [14], constituting a group of "outliers". Since a TSG is functioning in normal cells, it is expected to have Pol II in normal cells and belong to the outliers. Even if silencing of a TSG is a very rare event due to the Pol II binding, a cell with its silencing will become dominant due to the resultant growth advantage, and the TSG appears as if it is susceptible to methylation. Indeed, by searching outliers with high expression in normal cells, TSGs have been successfully identified [11,23,24]. However, some TSGs, including *CDKN1A*, are known to have no or low expression in normal cells and are activated by extracellular stimuli [25,26].

In this study, we aimed to show that a significant fraction of known TSGs silenced by aberrant DNA methylation are outliers. Then, we will show that a different set of TSGs could be identified by searching for outliers with high Pol II and without H3K27me3 (epigenome-based outlier approach), compared to those identified by searching for outliers with high expression (expression-based outlier approach).

2. Materials and methods

2.1. Cell lines and primary breast cancer samples

Human breast cancer cell lines (BT-20, BT-474, HCC38, HCC1428, HCC1937, Hs 578T, MCF7, MDA-MB-231, MDA-MB-436, MDA-MB-468, SK-BR-3, T-47D, and ZR-75-1), normal fetal human colon epithelial cell line (FHC), and 293TN cell line were purchased from the American Type Culture Collection (Rockville, MD). Normal human mammary epithelial cells (HMECs) were purchased from Cambrex (East Rutherford, NL). Forty primary breast cancer samples were obtained from patients who underwent mastectomy or breast conserving surgery with informed consent. Resected primary samples were frozen and stored at −80 °C until the extraction of genomic DNA. The analysis was approved by the Institutional Review Boards.

2.2. Chromatin immunoprecipitation (ChIP)

Eight μg of chromatin extracted from HMECs was incubated with 0.7 μg of antibodies against Pol II (ab5095; Abcam, Cambridge, UK) or H3K27me3 (07-449; Millipore, Billerica, MA) overnight at 4 °C. Then the immunocomplex was collected with Dynabeads Protein A (Invitrogen Dynal AS, Oslo, Norway). Recovered DNA was dissolved in 33 μl of 1 × TE (10 mM Tris-HCl and 1 mM EDTA) [14], and one μl of DNA was used for ChIP-quantitative PCR (ChIP-qPCR) using primer sets listed in Supplementary Table 1. The specificity of the ChIP assay was assessed by using primers for control regions whose histone modification statuses were already known [14] (Supplementary Fig. 1).

2.3. Analyses of microarray data

DNA methylation data of 9838 promoter CGIs in HMECs, BT-474, MCF7, MDA-MB-231, MDA-MB-468, and ZR-75-1 were obtained from our previous study using methylated DNA immunoprecipitation (MeDIP)-CGI microarray analysis [27]. Pol II binding and H3K27me3 data of 9838 promoter CGIs in HMECs were obtained from our previous study using ChIP-CGI microarray analysis [14]. Pol II binding and H3K27me3 levels of each CGI were evaluated using the average of the signal ratio [Cy5 signal (bound)/Cy3 signal (input)] of the probes within a NFR. CGIs with a mean signal ratio larger than that at the 80th percentile of the total probes were considered as CGIs with high Pol II binding or H3K27me3, and those with an average signal ratio smaller than that at the 20th percentile of the total probes were considered as CGIs with low Pol II binding or H3K27me3. Other CGIs were considered as genes with intermediate Pol II binding or H3K27me3 levels. Gene expression data in HMECs were obtained from our previous study using the GeneChip Human Genome U133 Plus 2.0 microarray (Affymetrix, Santa Clara, CA), and genes were classified into those with high, moderate, and low expression levels, as described previously [14].

2.4. Quantitative methylation-specific PCR (qMSP)

Genomic DNA was digested with *Bam*HI, and one μg of digested DNA was used for bisulfite modification [28]. qMSP was performed using primer sets specific to methylated target loci and to the *Alu* repeat sequence (Supplementary Table 2). Modified DNA was dissolved in 40 μl of 1 × TE, and one μl was used for qMSP. The number of DNA molecules in a sample was obtained by comparing its amplification curve with those of standard samples with known numbers of DNA molecules. DNA methylation levels were calculated as the percentage of the methylation reference (PMR) = [(number of molecules methylated at a target CGI in a sample)/(number of *Alu* repeat sequences in the sample)]/[(number of molecules methylated at the target CGI in a fully methylated DNA)/(number of *Alu* repeat sequences in the fully methylated DNA)] × 100 [29,30]. Genomic DNA treated with *SssI* methylase (New England Biolabs, Beverly, MA) was used as a fully methylated DNA. Since the copy number of the target CGI was normalized to the copy number of the *Alu* repeat sequence, PMR could reach more than 100% when the locus containing the target CGI had an increased copy number.

2.5. Quantitative reverse transcription-PCR (qRT-PCR)

Total RNA was isolated using ISOGEN (Nippon Gene, Tokyo, Japan). cDNA was synthesized from one μg of total RNA using SuperScript III reverse transcriptase and an oligo $(dT)_{12-18}$ primer (Life Technologies, Carlsbad, CA). qRT-PCR was performed using primer sets listed in Supplementary Table 3, as described previously [31]. The number of cDNA molecules in a sample was obtained by comparing its amplification curve with those of standard samples with known numbers of cDNA molecules, and the number of target cDNA molecules was normalized to those of human *GAPDH* cDNA molecules.

2.6. 5-Aza-dC and trichostatin A (TSA) treatment

T-47D cells were seeded on day 0, and exposed to a designated dose of freshly prepared 5-aza-dC (SIGMA-ALDRICH, St. Louis, MO) for 24 h on days 1 and 3, and treated with TSA (SIGMA-ALDRICH) on day 4. The cells were harvested on day 5, and total RNA was extracted.

2.7. Knockdown of DZIP1 and cell growth assay

DZIP1 was knocked down by two different short hairpin RNAs (shRNA1 and shRNA2; Supplementary Table 4) designed by using siDirect version 2.0 [32], and shRNA for firefly luciferase was used as the control. The sense and antisense oligonucleotides containing each shRNA sequence were annealed, and then ligated into *Bam*HI/*Eco*RI sites of pGreenPuro™ shRNA Cloning and Expression Lentivector (System Biosciences, Mountain View, CA). The shRNA expression construct was co-transfected with pPACK Packaging Plasmid Mix (System Biosciences) into 293TN cells using Lipofectamine™ with Plus™ Reagent (Life Technologies). Medium containing pseudovirus was collected 48 h after transfection, and stored at −80 °C until use. HCC1937 and MDA-MB-436 cells were infected with pseudovirus according to the manufacturer's protocol.

Cell growth was analyzed by seeding cells at a density of 5×10^4 cells/6 cm plate on day 0, and counting the cell numbers on days 1, 2, 3, 4, 5, and 8 using TC™ Automated Cell Counter (Bio-Rad Laboratories, Hercules, CA).

3. Results

3.1. Confirmation of known TSGs as outliers

To confirm that TSGs are outliers to the general rule of genes methylated in cancer cells, we first analyzed six known TSGs silenced in breast cancer by aberrant DNA methylation, *BRCA1*, *CDKN2A* (*p16*), *HOXA5*, *MASPIN*, *RASSF1A*, and *RBP1* [33–38]. The Pol II binding and H3K27me3 levels of these six genes were analyzed in HMECs (Fig. 1A). Among the six TSGs, *BRCA1*, *HOXA5*, and *MASPIN* had high Pol II binding and almost undetectable levels of H3K27me3 in the HMECs, demonstrating that these genes are outliers. In contrast, *RASSF1A* had low Pol II binding and high H3K27me3 levels, showing that *RASSF1A* was methylated according to the rule. *CDKN2A* and *RBP1* had intermediate levels of Pol II binding and H3K27me3. We also analyzed expression levels of the six TSGs in normal cells. Among the three outliers, *MASPIN* was expressed at high levels (active Pol II), and *BRCA1* and *HOXA5* were expressed at low levels (stalled Pol II) (Fig. 1C). In the case of three TSGs known to be methylation-silenced in colon cancer, *MLH1* and *RASSF1A* [7] had high levels of Pol II binding and low expression levels (stalled Pol II) (Fig. 1B and D). These results indicated that a significant fraction of TSGs are "outliers" to the general rule of genes methylated in cancer cells, and that the outliers could be classified into those with active Pol II and with stalled Pol II.

3.2. Screening of outliers in breast cancer based on their epigenome statuses

Confirming that three of six known TSGs in breast cancer and two of three known TSGs in colon cancer are outliers, we pro-