Oncogene (2017), 1–12

www.nature.com/onc

ORIGINAL ARTICLE

Epigenetic regulation of RNA polymerase III transcription in early breast tumorigenesis

J-L Park[1,2,12], Y-S Lee[3,12], M-J Song[4], S-H Hong[3], J-H Ahn[5], E-H Seo[1,2], S-P Shin[6], S-J Lee[6], BH Johnson[7], MR Stampfer[8], H-P Kim[4,9,10], S-Y Kim[1,2] and YS Lee[7,11]

RNA polymerase III (Pol III) transcribes medium-sized non-coding RNAs (collectively termed Pol III genes). Emerging diverse roles of Pol III genes suggest that individual Pol III genes are exquisitely regulated by transcription and epigenetic factors. Here we report global Pol III expression/methylation profiles and molecular mechanisms of Pol III regulation that have not been as extensively studied, using nc886 as a representative Pol III gene. In a human mammary epithelial cell system that recapitulates early breast tumorigenesis, the fraction of actively transcribed Pol III genes increases reaching a plateau during immortalization. Hyper-methylation of Pol III genes inhibits Pol III binding to DNA via inducing repressed chromatin and is a determinant for the Pol III repertoire. When Pol III genes are hypo-methylated, MYC amplifies their transcription, regardless of its recognition DNA motif. Thus, Pol III expression during tumorigenesis is delineated by methylation and magnified by MYC.

Oncogene advance online publication, 28 August 2017; doi:10.1038/onc.2017.285

INTRODUCTION

RNA polymerase III (Pol III) transcribes medium-sized ncRNAs and short interspersed elements such as Alu and MIR (mammalian-wide interspersed repeats). Many Pol III-transcribed ncRNAs (termed 'Pol III genes' hereafter), for example, 5S rRNA and transfer RNAs (tRNAs), have fundamental roles in cell growth. Early studies indicated that Pol III transcription is elevated in cancer, because the Pol III machinery is suppressed by tumor suppressors such as p53, Rb, PTEN or activated by the MYC oncogene (reviewed in White[1]). Among them, MYC is thought to have an extensive impact on the expression of Pol III genes in cancer, because it is the master transcription factor (TF) that amplifies the expression of a significant number ($>15\%$ of total) of genes when it is overexpressed in a number of malignancies (reviewed in Dang[2]). MYC activates classic Pol III genes 5S rRNA and tRNA.[3,4]

Epigenetic regulation, especially chromatin structure, has a critical role in the expression of Pol III genes (reviewed in Bhargava[5] and Park *et al.*[6]). Although CpG methylation is the most common epigenetic modification on DNA, silencing of Pol III genes by methylation has not been examined thoroughly in contrast to the abundant knowledge of Pol II genes (reviewed in Klose and Bird[7]). Although the transcription of tRNAs and U6 snRNA has been shown to be inhibited by methylation,[8,9] the biological significance of such suppression is questionable because these RNAs have indispensable cellular functions.

Besides fundamental biological roles of classic Pol III genes, their unconventional roles as well as new Pol III genes are emerging (reviewed in Hu *et al.*[10]). We have recently identified a Pol III gene, called nc886, whose silencing by CpG hyper-methylation is of prognostic significance in several malignancies.[11–14] nc886 is the first such case among Pol III genes. In addition to nc886, a significant fraction of Pol III genes is expected to be silenced in various normal and pathological conditions. For example, Alu and viral ncRNAs are shown to be suppressed by methylation[15–19] and their silencing is of biological significance in maintaining the stability of the human genome and as a viral defense mechanism. Another example is tRNA cleavage products that play gene-regulatory functions in a sequence-specific manner (reviewed in Fu *et al.*[20]). Thus, selective expression of a subset of tRNAs from a total of 608 tRNA-coding loci in the human genome could be crucial in certain biological contexts.

However, an extensive study on Pol III regulation has been hampered by some technical issues. Transcription units of many Pol III genes are identical or have highly similar sequences that are scattered throughout the genome in multiple copies. Therefore, it is challenging to correlate the transcription rate from a single genomic locus to the steady-state RNA level. For example, the regulation of individual Pol III transcription units by methylation and MYC has not been scrutinized, leaving important questions yet to be answered. Does MYC activate all Pol III genes or specifically activate the ones with the MYC recognition motif (E-box)? Is hypo-methylation a pre-requisite? How do methylation and/or MYC shape the Pol III gene expression pattern and how is it re-shaped during tumorigenesis?

For an in-depth study of molecular events during the human tumorigenesis, a cell-based model system is essential. A good

[1]Personalized Genomic Medicine Research Center, KRIBB, Daejeon, Korea; [2]Department of Functional Genomics, University of Science and Technology, Daejeon, Korea; [3]Rare Cancer Branch, Research Institute, National Cancer Center, Goyang-si, Korea; [4]Department of Environmental Medical Biology, Yonsei University College of Medicine, Seoul, Korea; [5]Department of Life and Nanopharmaceutical Sciences and Department of Oriental Pharmacy, Kyung Hee University, Seoul, Korea; [6]Immunotherapeutics Branch, Research Institute, National Cancer Center, Goyang-si, Korea; [7]Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, USA; [8]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA; [9]Institute of Tropical Medicine, Yonsei University College of Medicine, Seoul, Korea; [10]Brain Korea 21 Plus Project for Medical Science, Yonsei University College of Medicine, Seoul, Korea and [11]Department of Cancer Biomedical Science, Graduate School of Cancer Science and Policy, National Cancer Center, Goyang, Korea. Correspondence: Dr H-P Kim or Dr S-Y Kim or Dr YS Lee, Department of Cancer Biomedical Science, Graduate School of Cancer Science and Policy, National Cancer Center, 323 Ilsan-ro, Ilsandong-gu, Goyang, 10408, Republic of Korea.
E-mail: kimhp@yuhs.ac or kimsy@kribb.re.kr or yslee@ncc.re.kr
[12]Co-first authors.
Received 5 February 2017; revised 13 June 2017; accepted 3 July 2017

model should reflect *in vivo* events and should yield reproducible experimental data. Human tissue samples isolated *in vivo* have a limited proliferation capability, which imposes problems in obtaining enough quantities of experimental material for reliable results and precludes repeating experiments. *In vitro* cancer cell lines may have too many genomic abnormalities to provide data that reflect an *in vivo* situation. In addition, most cell lines lack genuine normal cell counterparts, and it is often difficult to discern which features are cancer-specific. Despite the limited availability of human-derived material, breast tissues are exceptional in that large quantities of normal tissues can be obtained through reduction mammoplasty and mastectomy surgeries and thus human mammary epithelial cell (HMEC) cultures could be established.

Primary cells isolated from normal breast tissues (184D in Figure 1a, 240L and 122L in Supplementary Figure S1A) have a finite life span and in order to become immortalized must overcome or bypass two tumor suppressive barriers: the 'stasis barrier' (=stress-associated senescence) and the 'immortalization barrier' (=replicative senescence) (reviewed in Stampfer *et al*.[21]). These two barriers can be overcome by various means: (1) Stasis can be overcome by errors in the Rb pathway, such as loss of $p16^{INK4a}$ expression or overexpressed cyclin D1, which can be accomplished by treatment with a chemical mutagen like benzo(a)pyrene (BaP), transduction of shRNA to p16 or of a cyclin D1:cdk2 construct, or errors generated by inactivation of the p53 tumor suppressor or dysregulated MYC expression. (2) Replicative senescence can be overcome by expression of sufficient telomerase activity to maintain telomere length, which can be accomplished by transduction of the MYC oncogene or unknown errors generated by telomere dysfunction. In some cases, these treatments can yield immortalized cell lines with additionally acquired malignant properties, such as anchorage-independent growth or tumorigenicity in immunosuppressed mice; anchorage-independent growth can also be conferred to non-malignant immortal lines by transduction of the mutated ErbB2 oncogene (Figure 1a).[22–27] Primary, immortalized and malignant cells in this study are designated respectively in green, purple and red letters; these color codes will be used throughout all figures. Having developed from normal tissues, HMEC cultures are an ideal *in vitro* cell line model recapitulating early events of breast tumorigenesis[21] (see also http://hmec.lbl.gov/mock/history.html).

Herein, as a first step to elucidate Pol III regulation in cancer, we have investigated the Pol III transcription and methylation in a HMEC culture system.[21] We have also studied the molecular mechanism as to how methylation and MYC regulate Pol III transcription, by using nc886 as a representative Pol III gene.

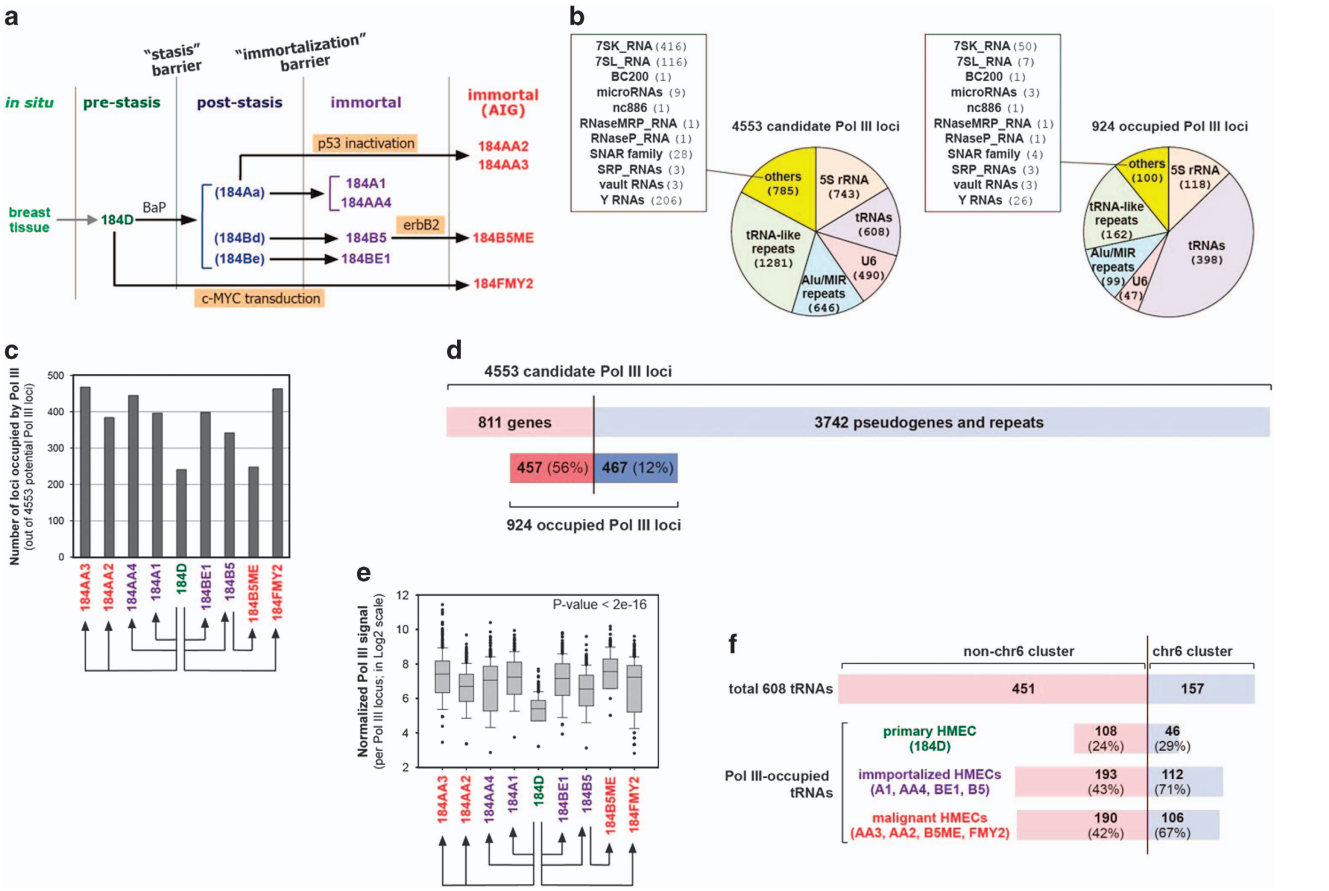


Figure 1. Pol III transcriptome in nine HMEC cultures. (**a**) Diagram illustrating the HMEC progression series derived from reduction mammoplasty specimen 184. (**b**) A pie chart of 4553 candidate Pol III genes (left panel) and 924 that were occupied by Pol III in at least one HMEC culture in our POLR3A ChIP-seq data (right panel). The full information on individual Pol III genes is in Supplementary Table S1 and S2. (**c**) Number of POR3A-occupied Pol III loci in the nine HMEC cultures derived from specimen 184. (**d**) A bar chart illustrating fractions of 'genuine Pol III genes' and 'pseudogenes and repeats'. See Supplementary Table S1 and S2 for full information. (**e**) A box plot of normalized Pol III occupancy per each locus in the nine 184-derived cultures. 10 000 000 of tags were used for normalization among cultures during the calculation of relative Pol III signal from POLR3A ChIP-seq reads. The analysis of variance (ANOVA) test was used to calculate a *P*-value. (**f**) A bar chart illustrating fractions of Pol III-occupied tRNA genes in a total of 608 tRNA loci (see Supplementary Table S3 for full information).

RESULTS

Landscape of global expression of Pol III genes (Pol III transcriptome) in HMEC cultures

To study epigenetic regulation of Pol III transcription during early tumorigenesis, we examined the Pol III transcriptome in nine HMEC cultures of the isogenic progression series derived from normal finite specimen 184 HMEC. RNA-sequencing is not an adequate method for Pol III genes, because their repetitive nature hampers the mapping of seq reads to genomic loci. Therefore, *Ch*romatin *I*mmuno*P*recipitation followed by high-throughput *seq*uencing (ChIP-seq) is commonly used as it captures flanking genomic sequences, which provide a unique sequence tag for mapping. We employed ChIP-seq of POLR3A, a Pol III catalytic subunit, as a proxy marker for Pol III transcription.

For ChIP-seq data analysis, we curated a list of 4553 candidate Pol III loci (Supplementary Table S1) by assembling a list from a previous report[28] and the tRNA database (http://gtrnadb.ucsc.edu/). These loci included known Pol III genes as well as the Alu and MIR repeat elements (Figure 1b). Some Pol III genes (BC200, nc886, vault RNAs (vtRNAs), RNaseMRP_RNA, RNaseP_RNA and SRP_RNA) are encoded at unique or a few loci, whereas others (5S rRNA, tRNAs, U6 snRNA, 7SK_RNAs, 7SL_RNAs and Y_RNAs) are present at >100 loci. In the latter case, only a minor fraction of them can produce full length transcripts (= genuine Pol III genes), whereas a majority of them are pseudogenes. Although these pseudogenes could be regarded as repeat elements, we followed their original annotation and did not classify them separately as a repeat element in the pie chart (Figure 1b), except for tRNAs. We regarded 608 loci in the tRNA database to be genuine tRNA genes and distinguished them from the other tRNA-like sequences that were classified as a repeat element (Figure 1b). During data analysis, we paid special attention to tRNA genes, because they are well defined and curated as compared with other Pol III genes.

From the raw ChIP-seq reads, we considered POLR3A peaks above a cutoff threshold (Poisson *P*-value $<1.00e-04$ and more than fourfold enrichment relative to input DNAs) to be significant and mapped them at the 4553 candidate Pol III loci. The validity of our ChIP-seq experiments and data processing were assured by several facts. First, the number of Pol III peaks mapped at the 4553 loci ranged from 241 to 468 in the nine 184-derived HMEC cultures (Figure 1c) and these numbers were comparable to previously reported numbers.[28–33] Second, when we analyzed the 924 loci that were Pol III-occupied in at least one HMEC (Supplementary Table S2), most of the genuine Pol III genes were included, whereas pseudogenes and repeats were largely excluded (Figure 1d). Third, the peaks were present at anticipated correct positions, as shown for nc886 and three vtRNAs as representative examples (Supplementary Figures S2A–D). Fourth, POLR3A ChIP-seq signals of nc886, a Pol III gene to be scrutinized later in this study, across the nine HMEC cultures were consistent with our Northern or quantitative reverse transcriptase-PCR data (Supplementary Figure S3; compare Figures 3a and b to be shown later).

The extent of Pol III occupancy, as indicated by the number of occupied Pol III loci (Figure 1c) and the relative Pol III signal estimated from seq reads (Figure 1e), was lower in normal primary 184D than any other HMEC lines, in agreement with the elevated Pol III activity during tumorigenesis in previous studies (reviewed in White[1]). This tendency was also seen in tRNA expression; 154 tRNA loci were bound by Pol III in 184D and this number nearly doubled in immortalized and malignant HMEC cultures (Figure 1f and Supplementary Figure S4A).

We further examined tRNA data. The extent of Pol III occupancy was increased during immortalization, but there was no further increase following malignant transformation (Figure 1f and Supplementary Figures S4A–B; compare HMEC cultures in purple and red letters). A feature of the genomic organization of human tRNA genes is the chr6 cluster (chr6: 26286753-28956860 in hg19) where 157 tRNA loci are housed within a region spanning about 2.5 mega-base pair. A majority of tRNA genes in this cluster were utilized in immortalized and malignant HMEC cultures (Figure 1f). According to the Pol III occupancy, we classified tRNAs (Supplementary Figure S4C and Supplementary Table S3); 78 tRNAs expressed in all nine HMEC cultures, 207 tRNAs in $1\sim7$ cultures and 210 tRNAs in none of them. The 78 constitutive tRNAs included a minimum set of 31 tRNAs necessary for decoding 61 codons to all 20 amino acids. Collectively from our data, the tRNA expression during early tumorigenesis was construed as follows: Primary cells express a set of tRNAs, small yet enough for the decoding role. As cells become immortalized, more tRNA genes are transcribed from the chr6 cluster and also other loci. No further increase seems to occur after immortalization. Some tRNAs are specifically expressed in individual HMEC cultures, and we speculate that those might play additional roles such as a gene-regulatory function upon their cleavage.

The global CpG methylation profile (methylome) and its impact on the Pol III transcriptome in HMEC

To compare expression data with methylation data, we obtained the methylome from the nine 184-derived HMEC cultures. Out of a total of 866 895 CpG sites in the methylation array, 3169 were localized in the 4553 Pol III loci (and ±300 bp). When we examined in the 924 loci that were actually bound by Pol III (in at least one HMEC line), 1421 CpG sites were mapped in 489 loci, with some individual Pol III loci having multiple CpG sites (Supplementary Table S4). We sorted the 1421 CpG sites into four groups according to their methylation values (an average of the nine HMEC cultures), compared with the POLR3A ChIP-seq signal, and found that higher methylation was associated with lower Pol III binding (Figure 2a).

This inverse correlation was also seen in a scatter plot between Pol III binding and methylation (an average methylation value if a locus has multiple CpG sites) (Figure 2b and Supplementary Table S5). In this chart, we noticed that data points were enriched in three areas (colored boxes in Figure 2b), except for few loci that were bound by Pol III despite hyper-methylation (designated by red color in Figure 2b) and identified to be some tRNA genes. Although these outliers are intriguing, we focused on the general propensity in interpreting data. The bulk of transcriptionally active Pol III loci were hypo-methylated (red-boxed) and were mostly composed of genuine Pol III genes. Inactive loci were hyper- or hypo-methylated (green- and blue-boxed, respectively) and contained most of the repeats and pseudogenes (Figure 2c). Inactive loci despite hypo-methylation could be partly explained by defective promoter elements in most pseudogenes. Also, the scatter plot of the repeats and pseudogenes indicated that most loci were inactive regardless of their methylation status, in contrast to the significant anti-correlation (Pearson's $r=-0.6795$) of genuine Pol III genes (Figure 2d). We further split the data into the six Pol III categories (see Figure 1b), plotted each of them and confirmed significant anti-correlation in tRNAs, U6 snRNA and others (Supplementary Figures S5A–C). In contrast, such anti-correlation was not seen in 5S rRNA, Alu and MIR repeats, and tRNA-like repeats (Supplementary Figures S5D and F). In the case of 5S rRNA, only two loci were genuine transcripts, but the remaining 23 loci were pseudogenes. We scrutinized these two loci and found them to be expressed in 184AA3 and 184FMY2 lines, but to be constitutively hyper-methylated even in these HMEC cultures (Supplementary Figure S6A–B). So, the expression of 5S rRNA could not be attributed to methylation and our data support a classic report in which 5S rRNA expression is not inhibited by methylation.[8] Of course, inspection of many other loci would be needed to make a concrete conclusion. Our data also corroborated previous papers having reported inhibition of tRNAs

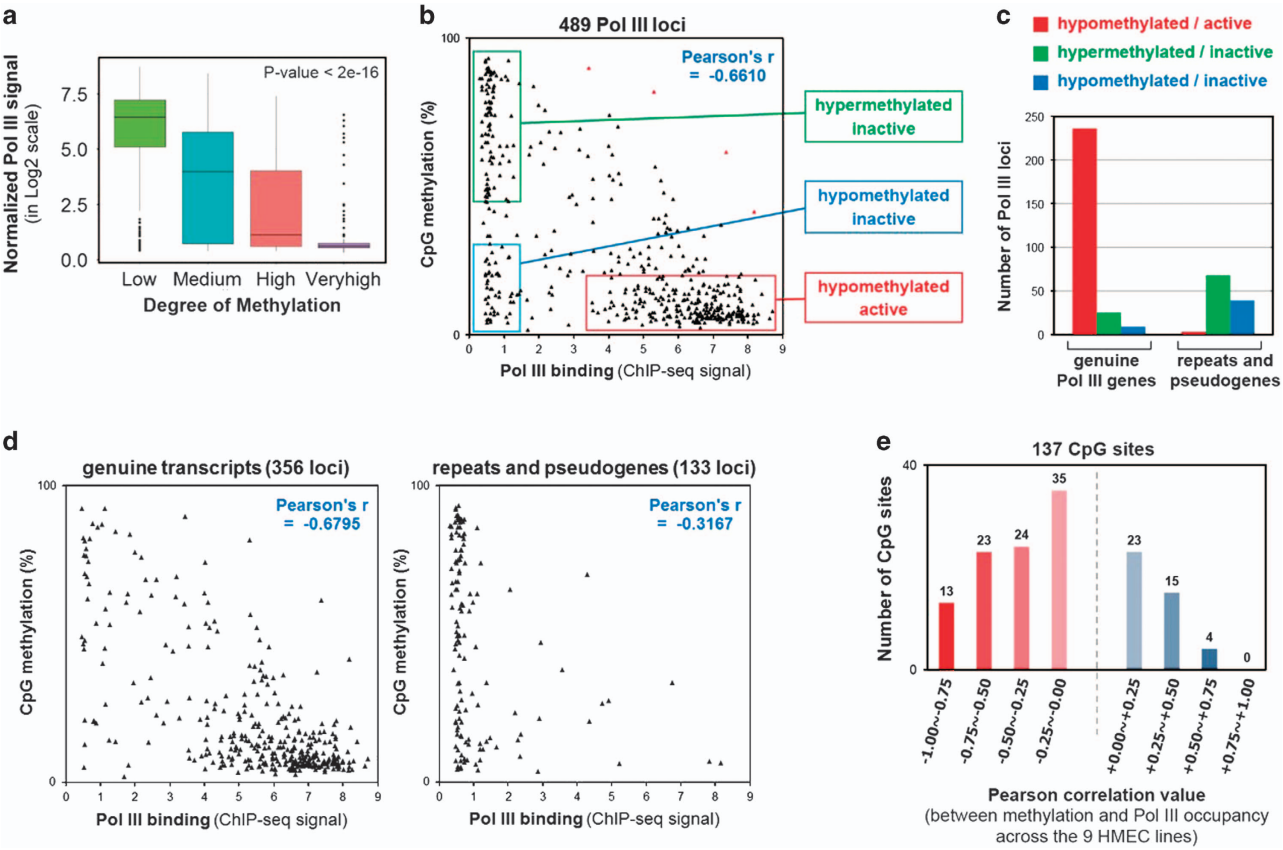


Figure 2. CpG methylome and its relation to Pol III transcriptome in the 184-derived progression series. (**a**) A box plot of normalized POLR3A ChIP-seq signals (calculated in Figure 1e) in each group of CpG sites. 1421 CpG sites in the 489 POLR3A-bound loci (occupied in at least one HMEC culture) were sorted according to methylation values (average-β) and were categorized in 'Low' ($\beta < 0.25$), 'Medium' ($0.25 \leqslant \beta < 0.50$), 'High' ($0.50 \leqslant \beta < 0.75$) and 'Veryhigh' ($0.75 \leqslant \beta$). The ANOVA test was used to calculate a *P*-value. See Supplementary Table S4 for full information. (**b**) A scatter plot between Pol III binding and methylation of 489 loci (listed in Supplementary Table S5). If a locus had multiple CpG sites, an average was calculated and used for plotting. From the distribution of data points, three hot areas were recognized and boxed. Four outlier data points are red-highlighted and are tRNAGly-GGG (chr1), tRNACys-GCA (chr7), tRNAGlu-CTC (chr6) and tRNACys-GCA (chr7). (**c**) Counts of Pol III loci ('genuine Pol III genes' versus 'repeats and pseudogenes') in the three groups of data points in (**b**). (**d**) Scatter plots of indicated categorized Pol III genes, drawn as described in (**b**). (**e**) A shortlist of 137 CpG sites was obtained from the 1421 sites (see the main text for details). Per each site, its methylation values and POLR3A ChIP-seq signal values were compared across the nine HMEC cultures to calculate a Pearson correlation value. A total of 137 correlation values were sorted and categorized, and then the number of CpG sites in each category (shown on the top of each bar) was counted and plotted.

and U6 snRNA by methylation.[8,9] Collectively, we proved the inhibition of genuine Pol III genes by methylation. In the case of repeats and pseudogenes, we surmise that their hyper-methylation was coincidental but not causative for their defective Pol III binding. So, our result was supportive to a recent paper that showed no effect of methylation on silencing of short interspersed element repeat elements.[19]

Pol III genes with nearby Pol II genes tended to be active (Supplementary Figure S7A), in agreement with a number of reports.[19,28–30,32–34] This correlation could be explained by methylation. CpG islands, blocks of sequence with a high frequency of CpG dinucleotides, are usually hypo-methylated and located near Pol II genes, especially house-keeping genes (reviewed in Deaton and Bird[35]). Also in our data, Pol III genes around a CpG island as well as the neighboring shore/shelf (located < 4 kb from an island) had an inclination to be located in proximity of Pol II genes (Supplementary Figure S7B) and to be hypo-methylated (Supplementary Figures S7C and D), as compared with the open sea (located > 4 kb away from an island). Taken together with our earlier data having shown the anti-correlation between methylation and expression of Pol III genes (Figures 2a and b), this analysis suggested that the expression of Pol III genes is dictated by the distribution of CpG islands and their methylation status, rather than the transcription *per se* of nearby Pol II genes.

Thus far we had used average values of the nine HMEC cultures to elucidate an overall relation between methylation and expression. Next, we analyzed individual values to cross-compare between the different HMEC cultures. To reasonably calculate a Pearson correlation between methylation and expression across the nine HMEC cultures, we made a shortlist of 137 CpG sites whose nine methylation values had a standard deviation $> 15\%$ and corresponding Pol III locus had a Pol III signal in $\geqslant 3$ HMEC lines. The 137 Pearson's *r*-values were significantly inclined to be negative (Figure 2e), ascertaining that methylation inhibits Pol III binding. Pol III genes with lowest Pearson's *r*-values (hence most negatively correlated) were nc886, vtRNA1-2 and some tRNA genes (shown in Supplementary Figure S8A–H, except for nc886, to be elaborated later), indicating that these genes were regulated by a dynamic change in methylation. During the course of tumorigenesis, global hypo-methylation and local hyper-methylation both occur, resulting in cancer genomes usually depleted of methylated CpGs on average (reviewed in Ehrlich[36]). However, we failed to see such a propensity along the

HMEC lineage, probably because the HMEC system represented the window of early tumorigenesis, while other reports looked into far advanced tumors obtained from clinical samples.

nc886 is an excellent model for Pol III study

Our Pol III transcriptome and methylome analysis indicated that methylation inhibits Pol III gene expression, and we wanted to substantiate the molecular mechanism. As stated in the Introduction, the repetitive nature of many Pol III genes imposes a technical problem in studying a single transcription unit. Furthermore, 5S rRNA and tRNAs are highly abundant and stable RNAs (Supplementary Figures S9A and B), exacerbating the problem. Actually, although our POLR3A ChIP-seq data indicated Pol III binding at certain confined loci in certain specific HMEC cultures, this variability was obscured when we observed the constant steady-state RNA level of 5S rRNA (Figures 3a and b and Supplementary Figure S10). To avoid these problems, in many Pol III studies precursor tRNAs (pre-tRNAs) are measured. However, it is questionable whether quantitative reverse transcriptase-PCR of pre-tRNAs genuinely measures the transcriptional rate, because they are processing intermediates whose steady-state levels are very low as determined both by a transcription rate and a maturation rate. This caveat explains why the expression of several pre-tRNAs did not show any recognizable pattern in the HMEC cultures (Figure 3b and Supplementary Figure S10).

nc886 is a recently identified Pol III gene whose expression level is critical in determining cell death/proliferation by controlling the activity of Protein Kinase R.[37] nc886 expression is silenced in several malignancies, indicating a tumor suppressor role therein.[11–14] On the other hand, nc886 plays an oncogenic role in thyroid cancer.[38] Our data from ovarian cancer (YSL, unpublished data) indicate that nc886 suppresses proliferation of cancer cells whereas it promotes their metastatic potential. Between cell proliferation and metastasis, the tumorigenic role differs according to tumor stage and determines whether nc886 is a tumor suppressor or an oncogene in that context. Commensurate with these contradictory roles, nc886 expression in cancer has two opposite facets and is exquisitely regulated as will be shown below.

The utilization of nc886 can circumvent aforementioned problems as well as provide several advantages for Pol III study: (1) it is encoded at a single genomic locus, (2) the product RNA is not too stable (half-life = 71.5 min in Supplementary Figures S9A and B),[37] (3) the RNA expression level is high enough to be measured reliably by Northern hybridization, (4) nc886 has a CpG island and is epigenetically silenced, (5) nc886 is activated by MYC (see below). Collectively, nc886 is an excellent model to investigate Pol III regulation whose interesting aspects have been largely concealed until now.

We began the nc886 study by characterizing it as to Pol III gene type, because each type has a distinct promoter architecture and Pol III subunit assembly and importantly is differently regulated by methylation (see Supplementary Figure S5). Pol III promoters are classified into three types; type I (for 5S rRNA), type II (for tRNAs) and type III (for U6 snRNA) (reviewed in Schramm and Hernandez,[39] see Supplementary Figure S11A). Previous ChIP-seq data showed that the nc886 locus is occupied by BRF1, but not by BRF2[28,31] (see Supplementary Figure S11B for summary). Because BRF2 is specific to type III (see Supplementary Figure

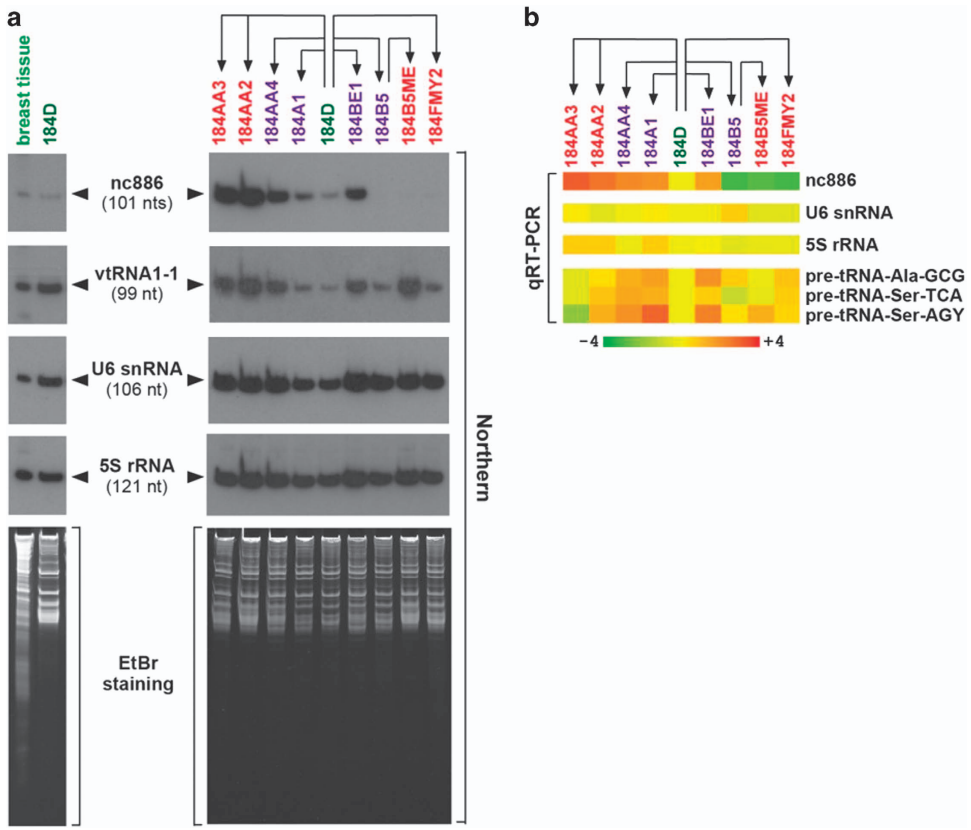


Figure 3. Expression of nc886 and some other Pol III genes in 184-derived progression series. (**a**) Northern hybridization of indicated Pol III genes. EtBr staining is shown for equal loading. The relationship among the different HMEC cultures in the progression series is depicted by lines and arrows on the top. (**b**) A heat map showing relative expression levels measured by qRT-PCR. Green and red colors indicate low and high expression relative to 184D (yellow), calculated from $2^{-\Delta Ct}$ values. The original ΔCt values are shown in Supplementary Figure S10.

S11A), nc886 is certainly not a type III. Type I is also highly improbable, because nc886 lacks an intermediate element that is seen specifically in the 5S rRNA gene (Supplementary Figure S11A and C). Like tRNA genes, nc886 has well conserved internal promoter elements box A and B. Also, nc886 is more similar to a tRNA gene than 5S rRNA, in terms of signal intensities and relative positions of three Pol III subunits; BRF1, BDP1 and POLR3D;[31] see Supplementary Figure S11B for summary). Collectively, nc886 is a type II Pol III gene.

nc886 expression in HMEC lines

nc886 expression among the HMEC cultures was of two opposite patterns, illustrative of its two contradictory roles in cancer as mentioned earlier. In one arm (the 184A-series lines shown on the left side of 184D in Figure 3a), nc886 expression was increased in the order of malignant > immortalized > primary cells (184D and normal breast tissue). This increase was also seen in other HMEC progression series (240L-derived and 122L-derived HMEC lines in Supplementary Figure S1A–C). Another Pol III gene, vtRNA1-1, was similarly increased (Figure 3a). These data are in agreement with previous literatures having reported that Pol III activity is elevated in cancer.[3,4,40–44] As compared with nc886, the expression of U6 snRNA was increased, but only moderately. A total of 5S rRNA levels were relatively constitutive (Figures 3a and b and Supplementary Figure S10). The transcription from some 5S rRNA loci was likely to be elevated based on our ChIP-seq data and also previous literatures;[3,4] however, this elevation appeared to be dominated by its abundance and stability, leading to a relatively constant steady-state RNA level.

In the other arm (the 184B-series lines shown on the right side of 184D in Figure 3a), nc886 expression was silenced in some lines (Figures 3a and b). These nc886-silenced lines (184B5, 184B5ME and 184FMY2) will be designated nc886$^-$ hereafter. Likewise, nc886-expressing cells will be designated nc886$^+$ (or nc886$^+$ to nc886^{+++++} when relative nc886 expression levels need to be indicated). Comparison of 184B5 (nc886$^-$) with other nc886$^+$ immortalized lines (184A1, 184AA4, 184BE1; see Figure 1a for lineage) suggested that a certain stochastic event during the immortalization stage led to nc886 silencing. Once this occurred in 184B5, the silencing was maintained in its progeny, 184B5ME.

nc886 silencing in HMEC lines by CpG DNA methylation

nc886 has a CpG island at the promoter region (Figure 4a) and its silencing by DNA hyper-methylation has been reported in several malignancies to be of clinical significance in patient prognosis.[11–14] nc886 silencing was also seen in some HMEC cultures at the immortalization stage (Figures 3a and b). Of note, a switch in the CpG methylation pattern has been shown to occur at this stage.[45] We measured CpG methylation at the nc886 region by pyrosequencing and bisulfite sequencing to prove hyper-methylation in nc886$^-$ HMEC cultures (Figures 4b and c and Supplementary Figure S12). In addition, treatment of 5-Aza-2'deoxycytidine (AzadC), an inhibitor for DNA methyltransferases (DNMTs), led to re-expression of nc886 in nc886$^-$ lines (184FMY2 in Figure 4d and 184B5 to be shown later in Figure 5c). All of these data undoubtedly showed that hyper-methylation during an immortalization stage caused nc886 silencing in HMECs. The nc886 hyper-methylation could not be attributed to an alteration in DNMTs (DNMT1, DNMT3A and DNMT3B), because their expression levels and also overall methylation levels were not correlated with nc886 levels in the nine HMEC cultures (Supplementary Figure S13). Our analysis of breast cancer patients from The Cancer Genome Atlas revealed nc886 hyper-methylation in tumor samples relative to normal breast tissues (Figure 4e), indicating that nc886 methylation did occur in breast cancer patients and may be of clinical significance.

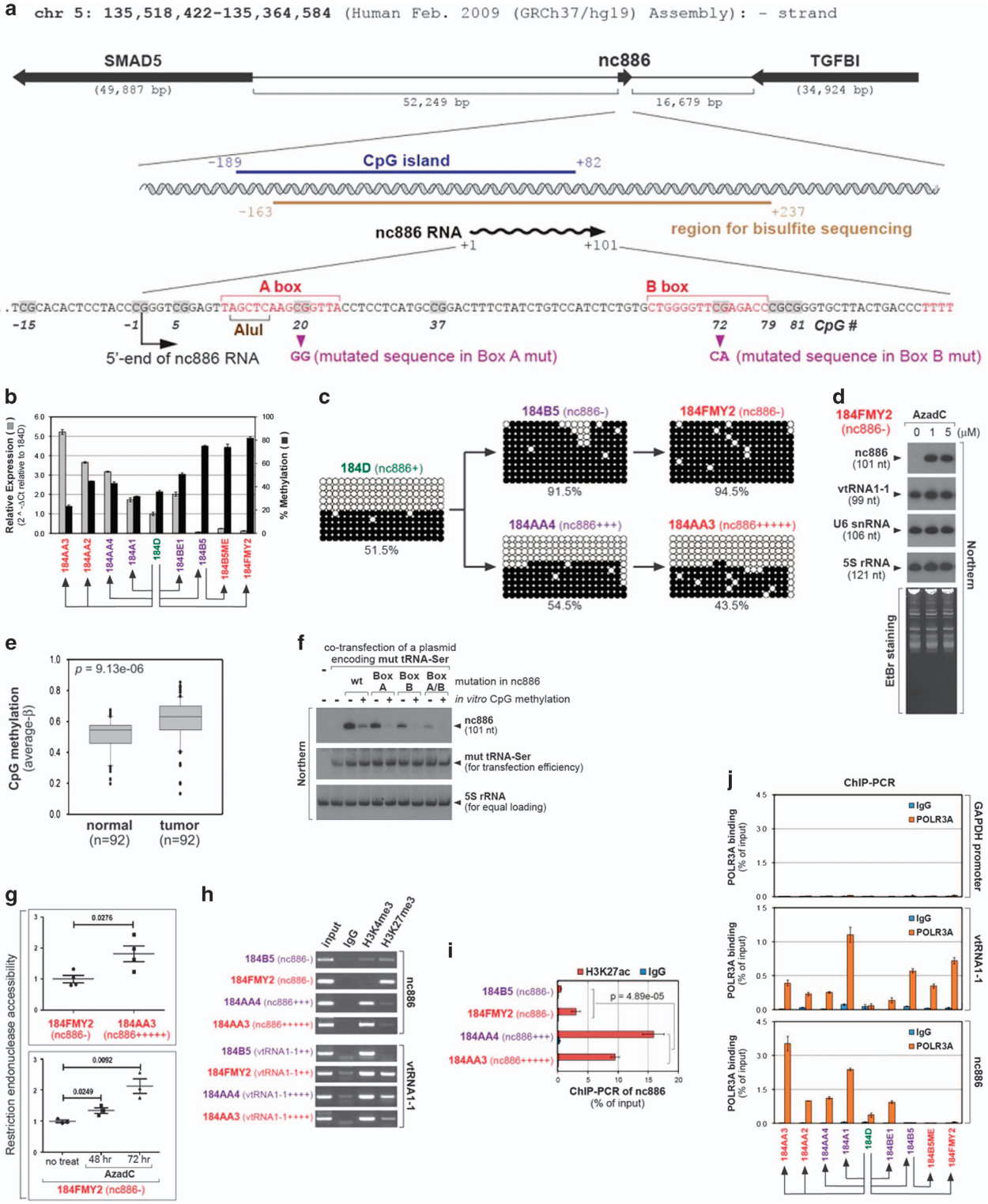
In addition to being a model Pol III gene for a mechanistic study, nc886 has several intriguing aspects in regard to methylation. nc886 is a metastable epiallele.[46] Also, nc886 shows allele-specific methylation,[14] with one allele of maternal origin being methylated in healthy normal individuals.[47] In agreement, our bisulfite sequencing data showed ~50% methylation in 184D which is closest lineage to the normal breast tissue. This means, from the oncology point of view, normal individuals inherently have loss of heterozygosity in nc886 methylation, and so aberrant hyper-methylation of the other allele is sufficient for its complete silencing. This notion, which is well illustrated in our bisulfite sequencing data of nc886$^-$ HMEC cultures relative to nc886$^+$ 184D (Figure 4c), explains why nc886 is so frequently silenced in tumors.[11–14] Another noteworthy point in Figure 4c is ~50% methylation in the 184AA4 and 184AA3 lines (nc886^{+++} and nc886^{++++++}) similar to 184D (nc886$^+$), indicating that the originally methylated allele was stably maintained during HMEC tumorigenesis and that the elevated nc886 expression was not owing to hypo- or de-methylation.

Next we interrogated how CpG methylation suppresses nc886. The consensus model, based on extensive studies on Pol II genes, is that CpG methylation leads to repressed chromatin and thereby induces silencing (reviewed in Fuks[48]). This model is presupposed to be true also for Pol III genes and is supported by a number of references as well as our data having shown the positive correlation of Pol III ChIP peaks with open chromatin and the expression of neighboring genes (Supplementary Figures S7A–D).[19,28–30,32–34] Nonetheless, it is also conceivable that CpG methylation *per se* might render the promoter less preferable for Pol III binding, and this alternative possibility has been supported by several pieces of indirect evidence. First, Pol III genes, especially type II, on a methylated plasmid (hence in a non-chromosomal context) were inefficiently transcribed when transfected into mammalian cells, injected into *Xenopus* oocytes, or added in *in vitro* transcription assays.[8,11,12,16,17,49] Second, TFIIIC, the Pol III subunit that recognizes type II promoter elements box A and B, did not bind to *in vitro* CpG-methylated DNA.[50]

As shown earlier (Supplementary Figure S11C), nc886 is a type II Pol III gene harboring box A and B, each of which contains a CpG dinucleotide sequence (Figure 4a). We reasoned that CpG methylation at box A and B might interfere with promoter recognition by Pol III and tested this hypothesis by nullifying the CpG residues via site-directed mutagenesis. Although we converted the CpG residues in a way that minimally harmed the box A and B consensus sequences (as referred in Oler *et al*.,[28] see Figure 4a for mutated sequences), nc886 was less efficiently transcribed in mutants than the wildtype (compare lanes 3, 5, 7 and 9 in Figure 4f). Importantly, when *in vitro* CpG-methylated, nc886 expression from the mutant DNA constructs was inhibited as much as the wild-type control (lane 3–4 versus 5–10 in Figure 4f and Supplementary Figure S14). Thus, these specific CpG residues are critical for transcription, but their methylation is dispensable for silencing. Taken together with previous reports,[17,49] it appears that a certain degree of CpG methylation at or near a Pol III gene, rather than specific residues, is sufficient for silencing.

We evaluated the effect of nc886 methylation on the chromatin structure by restriction enzyme accessibility assays and ChIP assays for methylated histones H3K4me3 and H3K27me3 as well as acetylated histones H3K27ac, H3Ac and H4ac. In Pol III and Pol II genes H3K4me3 and H3K27me3 are indicative of active and inactive promoters, respectively. In several ChIP-seq data both methylated histones have been detected as a sharp peak at the transcription start site of Pol III genes, whereas H3K27me3 peaks are broad in Pol II genes.[30] In general, H3Ac and H4Ac are markers for open chromatin status and therefore active transcription. Specifically, H3K27ac is enriched in the enhancer region of Pol II genes and is also associated with high transcriptional activity of Pol III genes (reviewed in Park *et al*.[6]). These two assays collectively

demonstrated open chromatin status of nc886^{+}/hypo-methylated lines (184AA3 and 184AA4) but repressed chromatin of nc886^{-}/hyper-methylated lines (184B5 and 184FMY2) (Figures 4g-i and Supplementary Figure S15). In addition, AzadC treatment in the 184FMY2 line, which released nc886 from its silencing, resulted in open chromatin status as shown by restriction enzyme accessibility assays (Figure 4g). The repressed chromatin of nc886^{-} cells did not allow Pol III binding to the nc886 region, as proven by the absence of nc886 in the POLR3A ChIP DNA (Figure 4j). In contrast, the POLR3A ChIP signal was evident in nc886^{+} HMEC cultures. Thus, hyper-methylation of nc886^{-} cells is associated with heterochromatin and lack of Pol III binding. The nc886 locus is unlikely to form long-range heterochromatin according to our data that its flanking genes TGFBI and SMAD5 were abundantly expressed even in nc886^{-} HMEC cultures (Supplementary Figure S16A) from a chromatin milieu distinct from that of nc886



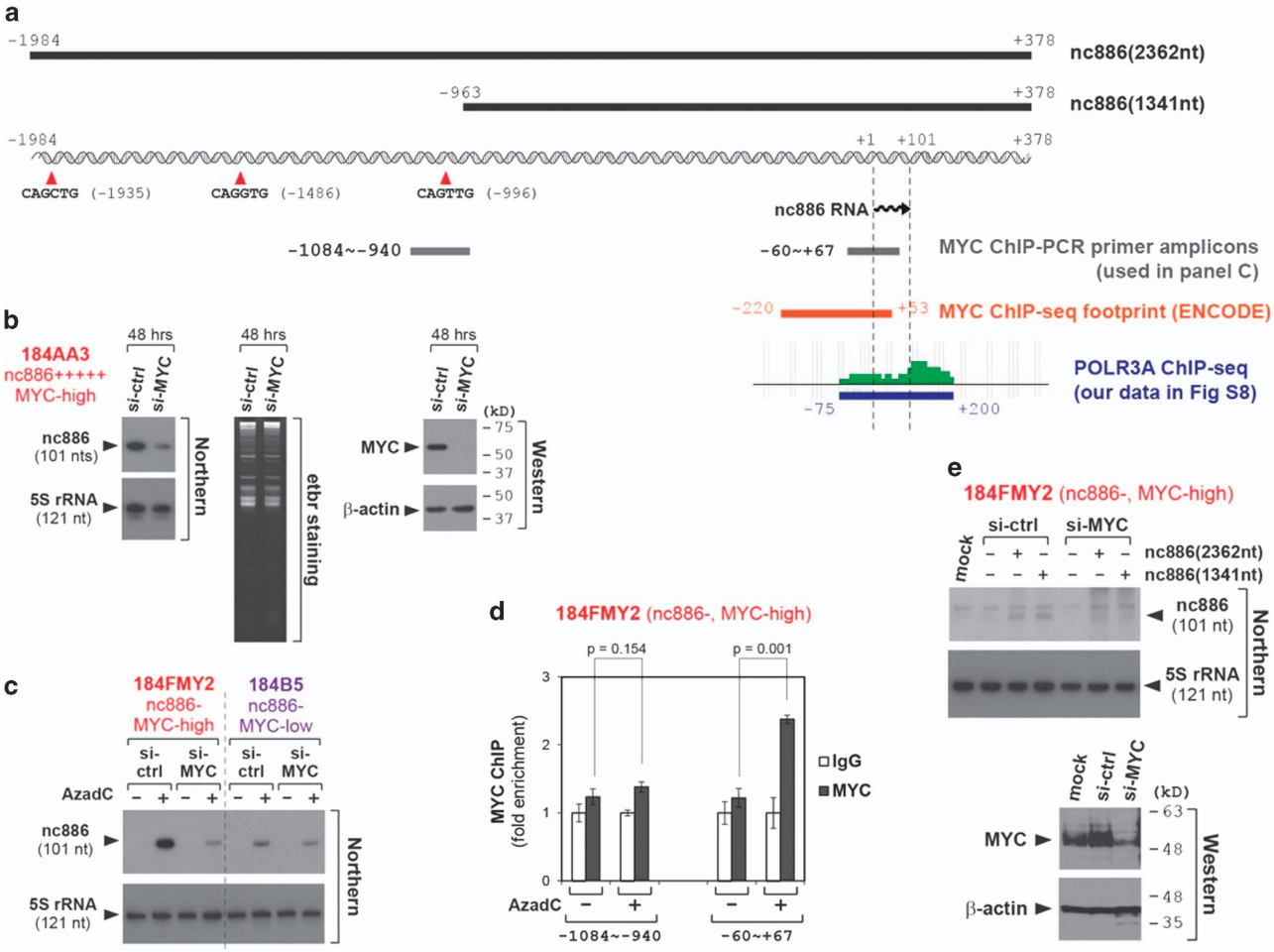


Figure 5. Amplification of nc886 expression by MYC independently of E-boxes. (**a**) A diagram illustrating the nc886 genomic region. Black bars on the top depict DNA constructs for a transfection experiment in **e**. E-boxes (5'-CANNTG-3') are indicated by red upward arrowheads. All nt coordinates are counted from the 5'-end of nc886 and all bars are drawn to the scale. (**b**) Northern hybridization of nc886 (and 5S rRNA) and western blot of MYC (and β-actin), after transfecting siRNA against MYC. EtBr staining is also shown for Northern because 5S rRNA is known to be affected by MYC. Protein molecular sizes from Xpert2 Prestained Protein Marker (GenDepot, Barker, TX, USA) are indicated on the right. (**c**) Northern hybridization of nc886 and 5S rRNA. AzadC treatment (at 2 μM) began at day 0, siRNA transfection was at day 2 and cells were harvested at day 4. (**d**) qPCR of MYC-ChIP. $2^{-\Delta\Delta Ct}$ values from ChIP/input DNA were calculated from MYC-ChIP and their enrichment relative to each IgG control is shown (*y* axis). The measurement was in triplicates. AzadC was treated at 2 μM for 4 days. (**e**) Northern hybridization after transfecting nc886-expressing DNA constructs in **a**. 184FMY2 cells were transfected with none ('mock'), non-targeting siRNA ('si-ctrl') and siRNA against MYC. After 2 days, cells were harvested for MYC Western (bottom panel) and split for DNA transfection. At 24 h post transfection of DNA, cells were harvested for Northern. The genuine nc886 band is indicated by arrowhead.

Figure 4. nc886 silencing by CpG DNA methylation. (**a**) A diagram depicting the nc886 genomic region. A wide nc886 locus spanning its flanking genes is on the top and two sequentially magnified views are on the middle and bottom. All nt coordinates are based on the 5'-end of nc886 as +1. (**b**) A bar graph of nc886 RNA expression and methylation by quantitative reverse transcriptase-PCR and pyrosequencing, respectively. RNA expression (reclaimed from Figure 3b and S10 data) is expressed in $2^{-\Delta Ct}$, with the value of 184D being set as 1. Percent methylation is an average of 4 CpG sites (#-15, −1, 5, 20 as shown in **a**; see Supplementary Figure S12 for methylation of individual sites). An average and the standard deviation from triplicates are shown. (**c**) Bisulfite sequencing results in indicated HMEC lines with their relative nc886 expression levels in parentheses. Ten clones per each CpG site were sequenced. Open and filled circles represent unmethylated and methylated CpG dinucleotides, respectively. (**d**) Northern hybridization of nc886 after 2 days of AzadC treatment. (**e**) A box plot of nc886 CpG methylation in 92 paired breast tumors and non-tumors tissues from the TCGA data set. The HumanMethylation 450 K BeadChip platform included 10 CpG sites (cg04481923, cg18678645, cg06536614, cg26328633, cg25340688, cg26896946, cg00124933, cg07845965, cg16615357 and cg18797653) at the nc886 locus. An average of the 10 CpGs was calculated and plotted. (**f**) Northern hybridization at 24 h post transfecting nc886-expressing plasmid 'pCR886(1.3k)' or its mutant derivatives at box A, B or both (designated as wt, Box A, BoxB and Box A/B, respectively; see also **a** for mutated sequences). Prior to transfection, these plasmid DNAs were mock-treated or CpG-methylated by the *M.Sss*I enzyme and proper methylation was assured by *Hha*I/*Hpa*II digestion of an aliquot (see Supplementary Figure S14 for details). The same transfection efficiency was assured by co-transfection of a plasmid expression mut tRNA-Ser (whose mutation was needed to distinguish from endogenous tRNA signal). (**g**) Restriction enzyme accessibility assays. nc886 Ct values in nested PCR were normalized to GAPDH and expressed in $2^{-\Delta\Delta Ct}$, with the value of untreated 184FMY2 line being set as 1. Higher values indicate more *Alu*I susceptibility owing to open chromatin structure. *P*-values are shown. (**h**) End-point PCR of nc886 and vtRNA1-1 (for comparison) on indicated ChIP DNA and input DNA. H3K4me3 and H3K27me3 are markers for open and repressed chromatin, respectively. (**i**) qPCR measurement of nc886 after H3K27ac (for open chromatin) ChIP. From $2^{-\Delta Ct}$ values from ChIP and input DNA, % bound was calculated (*x* axis). (**j**) qPCR measurement of indicated genes after POLR3A ChIP. % bound (*y* axis) was calculated as in **i**.

(Supplementary Figure S16B as compared with Figure 4h; Supplementary Figure S16C). This result is interesting when considering the CCCTC-binding factor, a barrier protein blocking the spread of heterochromatin (reviewed in Ong and Corces[51]). ENCODE ChIP-seq data indicate that there are two CCCTC-binding factor binding sites at each side of the nc886 region; one at ~2 kb upstream and the other at ~13 kb downstream of nc886. This suggests that the chromatin status of nc886 is insulated by CCCTC-binding factor, which would explain the lack of correlation between nc886 and TGFBI/SMAD5. It should be mentioned that these three genes might be co-regulated in other biological contexts although they were independently expressed in HMEC cultures. Overall, our data from nc886 are in good support of the conventional model that hyper-methylation induces heterochromatin and silencing.

Activation of nc886 expression in HMEC lines by MYC

Pol III transcription is elevated in cancer and so was it in our HMEC cultures. In the case of nc886, the elevated expression from 184D ($nc886^{+}$) to more tumorigenic cultures ($nc886^{+++}$ to $nc886^{+++++}$) cannot be attributed to hypo-methylation, because all those cells had the same 50% methylation. So, we looked into TFs, especially MYC, because it is a potent TF and cancer driver. In fact, several of the immortalized HMEC lines were generated with MYC transduction (Figure 1a and Supplementary Figure S1A).[23,24] The MYC expression level was the lowest in 184D (normal primary cells) but tended to be elevated in all other immortalized or further transformed HMEC cultures (Supplementary Figure S17). Also, ChIP-seq data from the Encyclopedia of DNA Elements (ENCODE) project suggested MYC binding to the nc886 region (Figure 5a). So, we examined MYC in nc886 regulation.

Knockdown (kd) of MYC decreased the nc886 expression level in the examined HMEC lines (Figures 5b and c) and also another breast cancer cell line, MDA-MB-231 ($nc886^{++++++}$, MYC-high;[37] Supplementary Figure S18A). 5S rRNA expression is also known to be activated by MYC, but was only marginally affected by MYC kd. This may be because 5S rRNA is too stable (see Supplementary Figures S9A and B) to elicit any significant reduction within 48 h. MYC kd resulted in less ChIP signal of POLR3A at nc886 and vtRNA1-1 (Supplementary Figure S18B), indicating that MYC is required for Pol III loading to promoters. We also did MYC kd in combination with AzadC treatment (Figure 5c). When 184FMY2 ($nc886^{-}$, MYC-high) was treated with AzadC, nc886 was re-expressed to a level comparable to 184AA3 cells ($nc886^{+++++}$, MYC-high) (Figures 4d and 5c). By comparison, nc886 was restored to a moderate expression level ($nc886^{+\sim++}$) by AzadC treatment in 184B5 cells ($nc886^{-}$, MYC-low) (compare lane 2 and 6 in Figure 5c). In both cell lines, MYC kd in this condition decreased nc886 to a basal expression level (compare lane 4 and 8 in Figure 5c) similarly to the level of 184D ($nc886^{+}$, MYC-lowest normal primary cells).

MYC binds to the recognition sequence E-box (5'-CACGTG-3'). Some TFs including MYC bind to distal loci and control transcription via forming a long-range chromatin loop (reviewed in Dang[2], and Kadauke and Blobel[52]). However, this possibility is unlikely for nc886 because chromatin interaction was not detectable at the nc886 locus when we examined *Ch*romatin *I*nteraction *A*nalysis *P*aired-*E*nd *T*ags data and Hi-C data from the ENCODE project (Supplementary Figures S19A and B). Based on this and also from the feasibility of our experiments using DNA constructs, we confined our search for an E-box within the proximal region (~2 kb) of nc886. TF prediction programs such as PROMO v3.0 using the TRANSFAC v8.3 database[53] and ConSite (http://consite.genereg.net/) did not detect any E-box therein. So, instead of searching for the strict E-box, we scanned the nc886 upstream region for any sequence matching to the E-box consensus sequence (5'-CANNTG-3') and were able to find three sites (Figure 5a). In our ChIP data, MYC binding was evident only in AzadC-treated 184FMY2 (Figure 5d) and was mapped near the POLR3A ChIP-peak harboring the nc886 transcript (Figures 5a and d) rather than in the ~1000 bp upstream region harboring an E-box consensus sequence. Our data suggested that the upstream E-boxes were dispensable, which was demonstrated by our transfection assays (Figure 5e). We prepared two different nc886-expressing plasmids, one with the E-boxes and the other without them, and transfected them into 184FMY2 cells. These two plasmids expressed comparable levels of nc886 (lane 3–4 in Figure 5e) and both of them were similarly inhibited upon MYC kd (lane 6–7 in Figure 5e).

Collectively, our MYC data were in good agreement with previous studies which show that the pre-requisite for MYC association to target promoters is an open chromatin status[54] and that MYC interacts with Pol III machinery rather than a DNA sequence element for stimulating Pol III transcription.[4] Although we prefer this simple model, it is also possible that MYC activates Pol III transcription via an epigenetic mechanism because MYC is known to recruit GCN5, a histone acetyltransferase, and its cofactor TRRAP to tRNA and 5S rRNA genes.[55] However, in our MYC kd experiments we did not observe any significant alteration of histone acetylation (Supplementary Figure S20) and could rule out the requirement of MYC for maintaining the acetylation. Nonetheless, a role of MYC for the establishment of open chromatin structure still remains to be a plausible scenario and waits for further investigation.

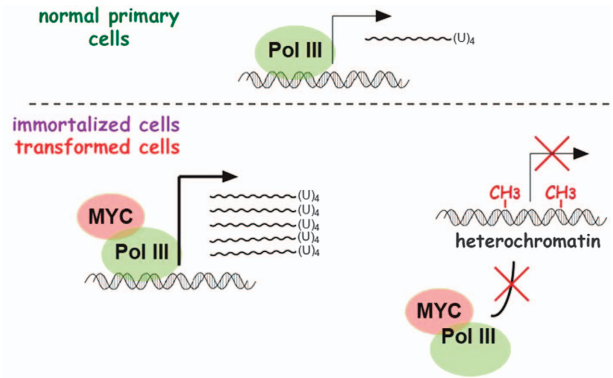


Figure 6. Summary cartoon illustrating the regulation of nc886 and other Pol III genes in the HMEC tumorigenesis model.

DISCUSSION

In summary, our data clarified several questions about the regulation of Pol III genes in cancer, with nc886 as a representative case (summarized in Figure 6). Normal finite cells sustain transcription of a subset of Pol III genes at a basal level while increasing the Pol III repertoire as they become immortalized. The basal level transcription of those Pol III genes is amplified by MYC in non-malignant and malignant immortalized lines, regardless of DNA sequence. Conversely in some other Pol III genes, hyper-methylation and the consequent heterochromatin formation precludes MYC and Pol III from binding to DNA. Based on our findings here, the landscape of the Pol III transcriptome is illustrated by methylation determining the framework and MYC boosting the color. In addition, we speculate that other various TFs add further details for exquisite shaping of the Pol III panorama in cancer. Our study is the first report to: (1) investigate the Pol III transcriptome and methylome in early tumorigenesis and (2) elucidate the molecular mechanism of MYC and epigenetic regulation in a single specific Pol III transcription unit.

This study only initiates understanding of Pol III regulation, which is far more complicated than previously thought. The three types of Pol III genes are regulated differently as shown here and the new information from nc886 in this study represents only type II Pol III genes. 5S rRNA genes (type I) cannot be explained by our simple model that hyper-methylation induces heterochromatin formation and gene silencing. We speculate that each of several thousands of Pol III genes are differently regulated and their transcriptional activity affects chromatin milieu in turn. All these questions need further investigation.

MATERIALS AND METHODS

Cell lines and other reagents

HMEC cultures were derived and grown as previously published.[23–27,56,57] A complete formulation of culture medium is available upon request. Sources of other reagents were described in Lee *et al*.[12] and Lee *et al*.[37]

Measurement and analysis of RNA

Northern hybridization and quantitative reverse transcriptase-PCR were performed as previously described in,[37] using primers summarized in Supplementary Table S6. Unless otherwise specified in the figure legend, all quantitative assays (quantitative reverse transcriptase-PCR, ChIP-PCR and so on) were done in biological triplicates for an average and the s.d., from which *P*-values were calculated by two-sided *t*-test.

Plasmid DNAs, siRNAs and transfection

'pCR886(1.3k)', a nc886-expressing plasmid, was used for transfection experiments (Figure 4f). To construct this plasmid, a 1355 nucleotide (nt) DNA fragment containing the nc886 genomic region (nt −977 ~ +378, referring to the 5'-end nt of nc886 RNA as +1) was PCR-amplified and then inserted into the pCR4-TOPO vector (Invitrogen, Carlsbad, CA, USA). Plasmids containing 2362 and 1341 nt of the nc886 region (see Figure 5a) were also made in the same manner. Since pCR4-TOPO lacks a mammalian promoter, nc886 expression from this plasmid is entirely driven by its own promoter. Plasmids carrying mutations at box A and/or B (see Figures 4a and f) were made from pCR886(1.3 k) by Quikchange site-directed mutagenesis. Prior to transfection, these plasmids were *in vitro* CpG-methylated by the *M.Sss*I enzyme, as described in Lee *et al*.[12] 'mut tRNA-Ser' (Figure 4f) was a mutant tRNA co-transfected to ascertain the same transfection efficiency. We cloned a PCR-fragment containing tRNA-Ser-TCA and its flanking genomic sequence (chr10:69524104-69524441) into pCR4-TOPO and then introduced a mutation at its anticodon loop (5′-AACT -3′ from the wild-type sequence 5′-TTGA-3′), so that ectopically expressed 'mut tRNA-Ser' could be measured without interference from an endogenous tRNA signal. All primers for PCR and mutagenesis are summarized in Supplementary Table S6. siRNAs against MYC was a Silencer Select siRNAs from Ambion (Thermo Fisher Scientific, Waltham, MA, USA) and its target sequence was 5′-gagcuaaaacggagcuuuu-3′. Plasmid DNAs and siRNA were transfected with Lipofectamine 2000 reagent and Lipofectamine RNAiMAX reagent (Invitrogen) respectively.

Measurement of CpG DNA methylation

Pyrosequencing of the nc886 promoter region was done as described in Lee *et al*.[12] The bisulfite sequencing procedure involved the following steps: conversion of genomic DNA by EZ DNA Methylation-Gold Kit (Zymo Research, Orange, CA, USA), PCR amplification of the nc886 region (nts −163 to +237; Figure 4a) from the bisulfite-modified DNA, cloning of the amplified DNA into pGEM-T Easy Vector (Promega, Southhampton, UK), and DNA sequencing of randomly chosen 10 white colonies. Primers for bisulfite sequencing were designed using MethPrimer (http://www.urogene.org/methprimer/index.html) and all primer information is in Supplementary Table S6.

Restriction endonuclease accessibility assays

Basically, this assay was to assess chromatin conformation by measuring sensitivity to *Alu*I restriction enzyme and the detailed procedure is described in Park *et al*.[58] The measured *Alu*I site is indicated in Figure 4a and the degree of digestion was measured by ligation-mediated PCR. All primers (adaptor primer, universal 5′-primers and gene-specific 3′-primers) are summarized in Supplementary Table S6. nc886 values were calculated from Ct values and were normalized to those of GAPDH.

ChIP

ChIP assays were performed as described previously[58] with minor modifications. In brief, cells were cross-linked with 1% formaldehyde, incubated in a swelling buffer (5 mM PIPES (pH 8.0), 85 mM KCl, 0.5% Nonidet P-40) and subjected to sonication in Buffer A (10 mM Tris-HCl (pH 8.0), 2 mM EDTA, 0.2% sodium dodecyl sulfate) using Bioruptor (Diagenode, Denville, NJ, USA). Chromatin samples were diluted in Buffer B (10 mM Tris-HCl (pH 8.0), 2% Triton X-100, 280 mM NaCl, 0.2% deoxycholate) and immunoprecipitated with antibodies against POLR3A (cat #ab96328 from Abcam, Cambridge, MA, USA), MYC (#9402 from Cell Signaling Technology, Danvers, MA, USA), H3K4me3 (#04-745 from EMD Millipore, Billerica, MA, USA), H3K27me3 (#07-449 from EMD Millipore) and H3K27Ac (#ab4729 from Abcam). Chromatin-antibody complexes were pulled down by Protein A/G Dynabeads (Invitrogen). After treating with proteinase K and reversing the cross-link, the amounts of immuneprecipitated DNA was measured by end-point PCR (for H3K4me3 and H3K27me3) or real-time PCR (for POLR3A, MYC and H3K27Ac) with indicated primers in Supplementary Table S6.

ChIP-seq

A total of 100–300 nt sized genomic libraries were generated with the input DNA and POLR3A ChIP DNA using the TruSeq ChIP Sample Prep Kit (Illumina, San Diego, CA, USA). In brief, input and ChIP DNA fragments were ligated to a pair of adaptors for sequencing, size-fractioned on a 2% agarose gel to isolate 200–300 nt fragments, and subjected to 18 cycles of PCR. Each library was diluted to 8 pM for 76 cycles of single-read sequencing on the Illumina NextSeq 500 per the manufacturer's recommended protocol. Seq reads were aligned on the human reference genome 19 with Bowtie2, and then the bam format file was converted to the sam format file using the SAM tools v1.4,[59] ChIP-peak calling and visualization were performed using a HOMER package using the default parameter setting.[60] Our ChIP-seq raw files are available from NCBI Sequence Read Archive (http://www.ncbi.nih.gov/sra/) via accession number SRP075723.

Measurement of CpG methylome

Infinium MethylationEPIC BeadChip (Illumina) was used for methylation array experiments per manufacturer's instruction. In brief, 500 ng of genomic DNA from nine HMEC cultures was treated with 20 µl sodium bisulfite solution that was included in the EZ DNA Methylation-Gold Kit (Zymo Research). Bisulfite-converted DNA (4 µl) was amplified using the Infinium Methylation Assay kit (Illumina). Amplified DNA was hybridized to Infinium MethylationEPIC BeadChip and scanned with the Illumina iSCAN system. CpG methylation values were calculated as average-β values using the GenomeStudio (V2010.3), per this equation: average-$\beta = M/(M+U+100)$, where M and U are methylated and unmethylated signal intensities, respectively. Measurements with P-value < 0.05 were considered to have a significant signal above background. All primary methylation array data were deposited in the GEO database under accession number GSE81939.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by a Research Scholar Grant, RSG-12-187-01—RMC from the American Cancer Society to YSL; grants NRF-2012M3A9D1054670 and NRF-2014M3C9A3068554 funded by the Ministry of Science, ICT and Future Planning and KRIBB Research Initiative to S-YK; a grant 2016R1A2B4014183 and 2017M3C9A5029978 funded by National Research Foundation of Korea to H-PK; National Cancer Center (Korea) intramural project # 1610090 to Y-SL; and US Department of Energy under Contract No. DE-AC02-05CH11231 to MRS.

REFERENCES

1 White RJ. RNA polymerase III transcription and cancer. *Oncogene* 2004; **23**: 3208–3216.
2 Dang CV. MYC on the path to cancer. *Cell* 2012; **149**: 22–35.

3 Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE *et al.* Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* 2012; **151**: 56–67.
4 Gomez-Roman N, Grandori C, Eisenman RN, White RJ. Direct activation of RNA polymerase III transcription by c-Myc. *Nature* 2003; **421**: 290–294.
5 Bhargava P. Epigenetic regulation of transcription by RNA polymerase III. *Biochim Biophys Acta* 2013; **1829**: 1015–1025.
6 Park JL, Lee YS, Kunkeaw N, Kim SY, Kim IH, Lee YS. Epigenetic regulation of noncoding RNA transcription by mammalian RNA polymerase III. *Epigenomics* 2017; **9**: 171–187.
7 Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 2006; **31**: 89–97.
8 Besser D, Gotz F, Schulze-Forster K, Wagner H, Kroger H, Simon D. DNA methylation inhibits transcription by RNA polymerase III of a tRNA gene, but not of a 5S rRNA gene. *FEBS Lett* 1990; **269**: 358–362.
9 Selvakumar T, Gjidoda A, Hovde SL, Henry RW. Regulation of human RNA polymerase III transcription by DNMT1 and DNMT3a DNA methyltransferases. *J Biol Chem* 2012; **287**: 7039–7050.
10 Hu S, Wu J, Chen L, Shan G. Signals from noncoding RNAs: unconventional roles for conventional pol III transcripts. *Int J Biochem Cell Biol* 2012; **44**: 1847–1851.
11 Lee HS, Lee K, Jang HJ, Lee GK, Park JL, Kim SY *et al.* Epigenetic silencing of the non-coding RNA nc886 provokes oncogenes during human esophageal tumorigenesis. *Oncotarget* 2014; **5**: 3472–3481.
12 Lee KS, Park JL, Lee K, Richardson LE, Johnson BH, Lee HS *et al.* nc886, a non-coding RNA of anti-proliferative role, is suppressed by CpG DNA methylation in human gastric cancer. *Oncotarget* 2014; **5**: 3944–3955.
13 Cao J, Song Y, Bi N, Shen J, Liu W, Fan J *et al.* DNA methylation-mediated repression of miR-886-3p predicts poor outcome of human small cell lung cancer. *Cancer Res* 2013; **73**: 3326–3335.
14 Treppendahl MB, Qiu X, Sogaard A, Yang X, Nandrup-Bus C, Hother C *et al.* Allelic methylation levels of the noncoding VTRNA2-1 located on chromosome 5q31.1 predict outcome in AML. *Blood* 2012; **119**: 206–216.
15 Liu WM, Maraia RJ, Rubin CM, Schmid CW. Alu transcripts: cytoplasmic localisation and regulation by DNA methylation. *Nucleic Acids Res* 1994; **22**: 1087–1095.
16 Banati F, Koroknai A, Salamon D, Takacs M, Minarovits-Kormuta S, Wolf H *et al.* CpG-methylation silences the activity of the RNA polymerase III transcribed EBER-1 promoter of Epstein-Barr virus. *FEBS Lett* 2008; **582**: 705–709.
17 Juttermann R, Hosokawa K, Kochanek S, Doerfler W. Adenovirus type 2 VAI RNA transcription by polymerase III is blocked by sequence-specific methylation. *J Virol* 1991; **65**: 1735–1742.
18 Xie H, Wang M, Bonaldo Mde F, Rajaram V, Stellpflug W, Smith C *et al.* Epigenomic analysis of Alu repeats in human ependymomas. *Proc Natl Acad Sci USA* 2010; **107**: 6952–6957.
19 Varshney D, Vavrova-Anderson J, Oler AJ, Cowling VH, Cairns BR, White RJ. SINE transcription by RNA polymerase III is suppressed by histone methylation but not by DNA methylation. *Nat Commun* 2015; **6**: 6569.
20 Fu Y, Lee I, Lee YS, Bao X. Small non-coding transfer RNA-derived RNA fragments (tRFs): their biogenesis, function and implication in human diseases. *Genomics Inform* 2015; **13**: 94–101.
21 Stampfer MR, LaBarge MA, Garbe JC. An integrated human mammary epithelial cell culture System for studying carcinogenesis and aging. In: Schatten H (ed). *Cell Mol Biol Breast Cancer*. Springer: NY, USA, 2013, pp 323–361.
22 Hines WC, Kuhn I, Thi K, Chu B, Stanford-Moore G, Sampayo R *et al.* 184AA3: a xenograft model of ER+ breast adenocarcinoma. *Breast Cancer Res Treat* 2016; **155**: 37–52.
23 Lee JK, Garbe JC, Vrba L, Miyano M, Futscher BW, Stampfer MR *et al.* Age and the means of bypassing stasis influence the intrinsic subtype of immortalized human mammary epithelial cells. *Front Cell Dev Biol* 2015; **3**: 13.
24 Garbe JC, Vrba L, Sputova K, Fuchs L, Novak P, Brothman AR *et al.* Immortalization of normal human mammary epithelial cells in two steps by direct targeting of senescence barriers does not require gross genomic alterations. *Cell Cycle* 2014; **13**: 3423–3435.
25 Garbe JC, Bhattacharya S, Merchant B, Bassett E, Swisshelm K, Feiler HS *et al.* Molecular distinctions between stasis and telomere attrition senescence barriers shown by long-term culture of normal human mammary epithelial cells. *Cancer Res* 2009; **69**: 7557–7568.
26 Stampfer MR, Garbe J, Nijjar T, Wigington D, Swisshelm K, Yaswen P. Loss of p53 function accelerates acquisition of telomerase activity in indefinite lifespan human mammary epithelial cell lines. *Oncogene* 2003; **22**: 5238–5251.
27 Stampfer MR, Bartley JC. Induction of transformation and continuous cell lines from normal human mammary epithelial cells after exposure to benzo[a]pyrene. *Proc Natl Acad Sci USA* 1985; **82**: 2394–2398.
28 Oler AJ, Alla RK, Roberts DN, Wong A, Hollenhorst PC, Chandler KJ *et al.* Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat Struct Mol Biol* 2010; **17**: 620–628.
29 Alla RK, Cairns BR. RNA polymerase III transcriptomes in human embryonic stem cells and induced pluripotent stem cells, and relationships with pluripotency transcription factors. *PLoS ONE* 2014; **9**: e85648.
30 Barski A, Chepelev I, Liko D, Cuddapah S, Fleming AB, Birch J *et al.* Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat Struct Mol Biol* 2010; **17**: 629–634.
31 Canella D, Praz V, Reina JH, Cousin P, Hernandez N. Defining the RNA polymerase III transcriptome: genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res* 2010; **20**: 710–721.
32 Carriere L, Graziani S, Alibert O, Ghavi-Helm Y, Boussouar F, Humbertclaude H *et al.* Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells. *Nucleic Acids Res* 2012; **40**: 270–283.
33 Moqtaderi Z, Wang J, Raha D, White RJ, Snyder M, Weng Z *et al.* Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat Struct Mol Biol* 2010; **17**: 635–640.
34 Canella D, Bernasconi D, Gilardi F, LeMartelot G, Migliavacca E, Praz V *et al.* A multiplicity of factors contributes to selective RNA polymerase III occupancy of a subset of RNA polymerase III genes in mouse liver. *Genome Res* 2012; **22**: 666–680.
35 Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes Dev* 2011; **25**: 1010–1022.
36 Ehrlich M. DNA methylation in cancer: too much, but also too little. *Oncogene* 2002; **21**: 5400–5413.
37 Lee K, Kunkeaw N, Jeon SH, Lee I, Johnson BH, Kang GY *et al.* Precursor miR-886, a novel noncoding RNA repressed in cancer, associates with PKR and modulates its activity. *RNA* 2011; **17**: 1076–1089.
38 Lee EK, Hong SH, Shin S, Lee HS, Lee JS, Park EJ *et al.* nc886, a non-coding RNA and suppressor of PKR, exerts an oncogenic function in thyroid cancer. *Oncotarget* 2016; **7**: 75000–75012.
39 Schramm L, Hernandez N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev* 2002; **16**: 2593–2620.
40 Felton-Edkins ZA, Kenneth NS, Brown TR, Daly NL, Gomez-Roman N, Grandori C *et al.* Direct regulation of RNA polymerase III transcription by RB, p53 and c-Myc. *Cell Cycle* 2003; **2**: 181–184.
41 Winter AG, Sourvinos G, Allison SJ, Tosh K, Scott PH, Spandidos DA *et al.* RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. *Proc Natl Acad Sci USA* 2000; **97**: 12619–12624.
42 Felton-Edkins ZA, White RJ. Multiple mechanisms contribute to the activation of RNA polymerase III transcription in cells transformed by papovaviruses. *J Biol Chem* 2002; **277**: 48182–48191.
43 Chen W, Bocker W, Brosius J, Tiedge H. Expression of neural BC200 RNA in human tumours. *J Pathol* 1997; **183**: 345–351.
44 Chen W, Heierhorst J, Brosius J, Tiedge H. Expression of neural BC1 RNA: induction in murine tumours. *Eur J Cancer* 1997; **33**: 288–292.
45 Novak P, Jensen TJ, Garbe JC, Stampfer MR, Futscher BW. Stepwise DNA methylation changes are linked to escape from defined proliferation barriers and mammary epithelial cell immortalization. *Cancer Res* 2009; **69**: 5251–5258.
46 Silver MJ, Kessler NJ, Hennig BJ, Dominguez-Salas P, Laritsky E, Baker MS *et al.* Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol* 2015; **16**: 118.
47 Romanelli V, Nakabayashi K, Vizoso M, Moran S, Iglesias-Platas I, Sugahara N *et al.* Variable maternal methylation overlapping the nc886/vtRNA2-1 locus is locked between hypermethylated repeats and is frequently altered in cancer. *Epigenetics* 2014; **9**: 783–790.
48 Fuks F. DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* 2005; **15**: 490–495.
49 Liu WM, Schmid CW. Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. *Nucleic Acids Res* 1993; **21**: 1351–1359.
50 Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 2010; **143**: 470–484.
51 Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet* 2014; **15**: 234–246.
52 Kadauke S, Blobel GA. Chromatin loops in gene regulation. *Biochim Biophys Acta* 2009; **1789**: 17–25.
53 Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* 2002; **18**: 333–334.
54 Guccione E, Martinato F, Finocchiaro G, Luzi L, Tizzoni L, Dall' Olio V *et al.* Myc-binding-site recognition in the human genome is determined by chromatin context. *Nat Cell Biol* 2006; **8**: 764–770.

55 Kenneth NS, Ramsbottom BA, Gomez-Roman N, Marshall L, Cole PA, White RJ. TRRAP and GCN5 are used by c-Myc to activate RNA polymerase III transcription. *Proc Natl Acad Sci USA* 2007; **104**: 14917–14922.
56 Labarge MA, Garbe JC, Stampfer MR. Processing of human reduction mammoplasty and mastectomy tissues for cell culture. *J Vis Exp* 2013; doi: 10.3791/50011.
57 Garbe JC, Holst CR, Bassett E, Tlsty T, Stampfer MR. Inactivation of p53 function in cultured human mammary epithelial cells turns the telomere-length dependent senescence barrier from agonescence into crisis. *Cell Cycle* 2007; **6**: 1927–1936.
58 Park K, Park JH, Yang WJ, Lee JJ, Song MJ, Kim HP. Transcriptional activation of the IL31 gene by NFAT and STAT6. *J Leukoc Biol* 2012; **91**: 245–257.
59 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
60 Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010; **38**: 576–589.

Supplementary Information accompanies this paper on the Oncogene website (http://www.nature.com/onc)