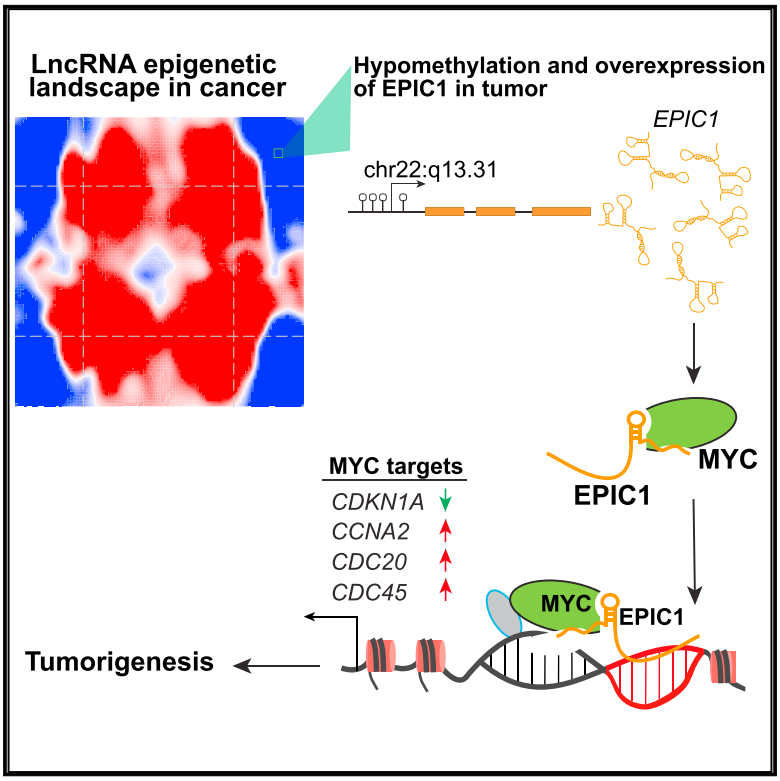
Article

Cancer Cell

lncRNA Epigenetic Landscape Analysis Identifies *EPIC1* as an Oncogenic lncRNA that Interacts with MYC and Promotes Cell-Cycle Progression in Cancer

Graphical Abstract



Highlights

- LncRNAs show a hypomethylation phenotype, in contrast to a CIMP phenotype in cancer
- *EPIC1* promotes breast tumorigenesis through regulating cancer cell-cycle progression
- *EPIC1* directly interacts with MYC protein through *EPIC1*'s 129–283 nt region
- *EPIC1* regulates MYC targets by enhancing MYC occupancy on its target promoters

Authors

Zehua Wang, Bo Yang, Min Zhang, ..., The Cancer Genome Atlas Research Network, Wen Xie, Da Yang

Correspondence

dyang@pitt.edu

In Brief

Wang et al. characterize the epigenetic landscape of lncRNAs genes across a large number of human tumors and cancer cell lines and observe recurrent hypomethylation of lncRNA genes, including *EPIC1*. *EPIC1* RNA promotes cell-cycle progression by interacting with MYC and enhancing its binding to target genes.

Wang et al., 2018, Cancer Cell *33*, 706–720
April 9, 2018
https://doi.org/10.1016/j.ccell.2018.03.006

CellPress

CellPress

Cancer Cell
Article

lncRNA Epigenetic Landscape Analysis Identifies *EPIC1* as an Oncogenic lncRNA that Interacts with MYC and Promotes Cell-Cycle Progression in Cancer

Zehua Wang,[1,4] Bo Yang,[1,4] Min Zhang,[1] Weiwei Guo,[1] Zhiyuan Wu,[1] Yue Wang,[1] Lin Jia,[1] Song Li,[1] The Cancer Genome Atlas Research Network, Wen Xie,[1] and Da Yang[1,2,3,5,*]
[1]Center for Pharmacogenetics, Department of Pharmaceutical Sciences, University of Pittsburgh, Pittsburgh, PA 15261, USA
[2]University of Pittsburgh Cancer Institute, University of Pittsburgh, Pittsburgh, PA 15261, USA
[3]Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA 15261, USA
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: dyang@pitt.edu
https://doi.org/10.1016/j.ccell.2018.03.006

SUMMARY

We characterized the epigenetic landscape of genes encoding long noncoding RNAs (lncRNAs) across 6,475 tumors and 455 cancer cell lines. In stark contrast to the CpG island hypermethylation phenotype in cancer, we observed a recurrent hypomethylation of 1,006 lncRNA genes in cancer, including *EPIC1* (epigenetically-induced lncRNA1). Overexpression of *EPIC1* is associated with poor prognosis in luminal B breast cancer patients and enhances tumor growth *in vitro* and *in vivo.* Mechanistically, *EPIC1* promotes cell-cycle progression by interacting with MYC through *EPIC1*'s 129–283 nt region. *EPIC1* knockdown reduces the occupancy of MYC to its target genes (e.g., *CDKN1A*, *CCNA2*, *CDC20*, and *CDC45*). MYC depletion abolishes *EPIC1*'s regulation of MYC target and luminal breast cancer tumorigenesis *in vitro* and *in vivo*.

INTRODUCTION

The most recent genome-wide characterization of the human cancer transcriptome has demonstrated that lncRNA expression is among the most pervasive transcriptional changes in cancer (Du et al., 2013; Iyer et al., 2015). Further experimental evidence indicates that lncRNAs can play important roles in tumorigenesis (Du et al., 2016; Prensner and Chinnaiyan, 2011; Schmitt and Chang, 2016; Zhu et al., 2016). Similar to protein-coding genes (PCGs), lncRNA expression is subject to changes in gene dosage (e.g., copy-number alterations) and promoter utilization (e.g., DNA methylation) that occur in cancer initiation and progression. In this regard, lncRNA genes can be targeted by cancer somatic alterations and thus play important roles in tumorigenesis. Recent studies focusing on the identification of copy-number alterations (Hu et al., 2014; Leucci et al., 2016; Yan et al., 2015) and cancer risk polymorphism in promoter regions (Guo et al., 2016) of lncRNA genes have provided evidence demonstrating that somatic/germline alterations of lncRNA in tumors can be "driver molecular events" leading to tumor initiation and progression.

Epigenetic regulation is one of the major mechanisms utilized to control lncRNA expression and tissue specificity (Amin et al., 2015; Guttman et al., 2009; Wu et al., 2010). Epigenetic alterations have been established as one of the hallmarks of tumorigenesis (Jones and Baylin, 2002; Shen and Laird, 2013). However, the epigenetic alterations of lncRNA genes and their consequences in cancer remain poorly characterized. Genome-scale studies have yielded important insights into DNA methylation changes in tumors (Irizarry et al., 2009; Noushmehr et al., 2010) but have mostly focused on PCG promoters.

Significance

Although global epigenetic alterations have been established as a prominent cancer hallmark, the epigenetic abnormality of lncRNA loci and their consequences in cancer development remain poorly characterized. We report an in-depth characterization of epigenetic landscape of lncRNA genes in 20 cancer types and discover that the expression of lncRNAs is recurrently epigenetically activated in tumors by hypomethylation. This study provides an integrative strategy of identifying lncRNA genes with oncogenic activity. Using this strategy, we have identified and validated *EPIC1* as an oncogenic lncRNA by interacting with the MYC protein and promoting cell-cycle progression. These discoveries expand upon the known mechanisms of MYC activation in cancer and pave the way to develop therapies that target MYC through its interaction with *EPIC1*.

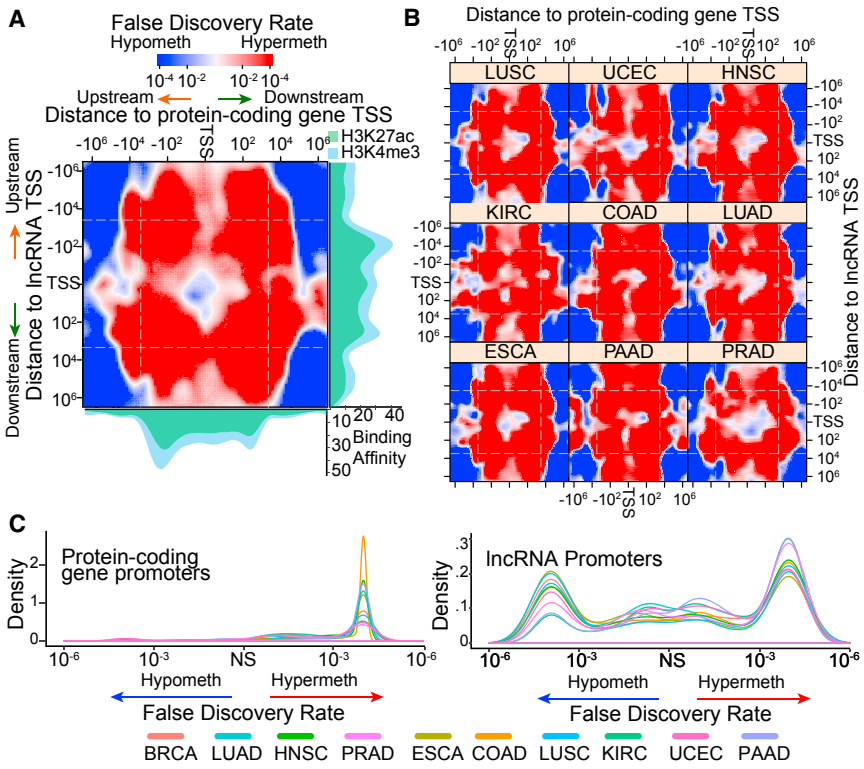


Figure 1. LncRNA and PCGs Have Distinct DNA Methylation Patterns in Ten Cancer Types

(A) Weighted density plot (kde2d.weighted [package: ggtern]) of differential DNA methylation (indicated by FDR values) of 100 windows within ±1,000 kb from transcription start sites (TSSs) are shown in breast cancer tissues. The windows are arranged based on their distances to PCG TSS (x axis) and lncRNA gene TSS (y axis). The promoter region is defined as ±3 kb (white dashed lines) from TSS. The hypermethylation region in tumor is shown as red, whereas the hypomethylation region is shown as blue. The average H3K27ac and H3K4me3 binding intensities are shown along with the x and y axes.

(B) Differential DNA methylation between tumors and matched normal tissues in nine cancer types.

(C) Distribution of the differential DNA methylation weighted density values (kde2d.weighted [package: ggtern]) within ±3 kb region (white dashed lines) of PCG TSS (left) and lncRNA TSS (right) in ten cancer types. NS, not significant.

See also Figure S1 and Table S1.

The efforts to characterize the lncRNA epigenetic landscape in cancers have labored under the limitations of an imperfect annotation of lncRNAs and a dearth of platforms that can detect lncRNA epigenetic and expression alterations in cancer. The emergence of large-scale cancer genomic/epigenetic projects, such as The Cancer Genome Atlas (TCGA) Research Network project, have provided an excellent opportunity to characterize the lncRNA epigenetic landscape in cancer.

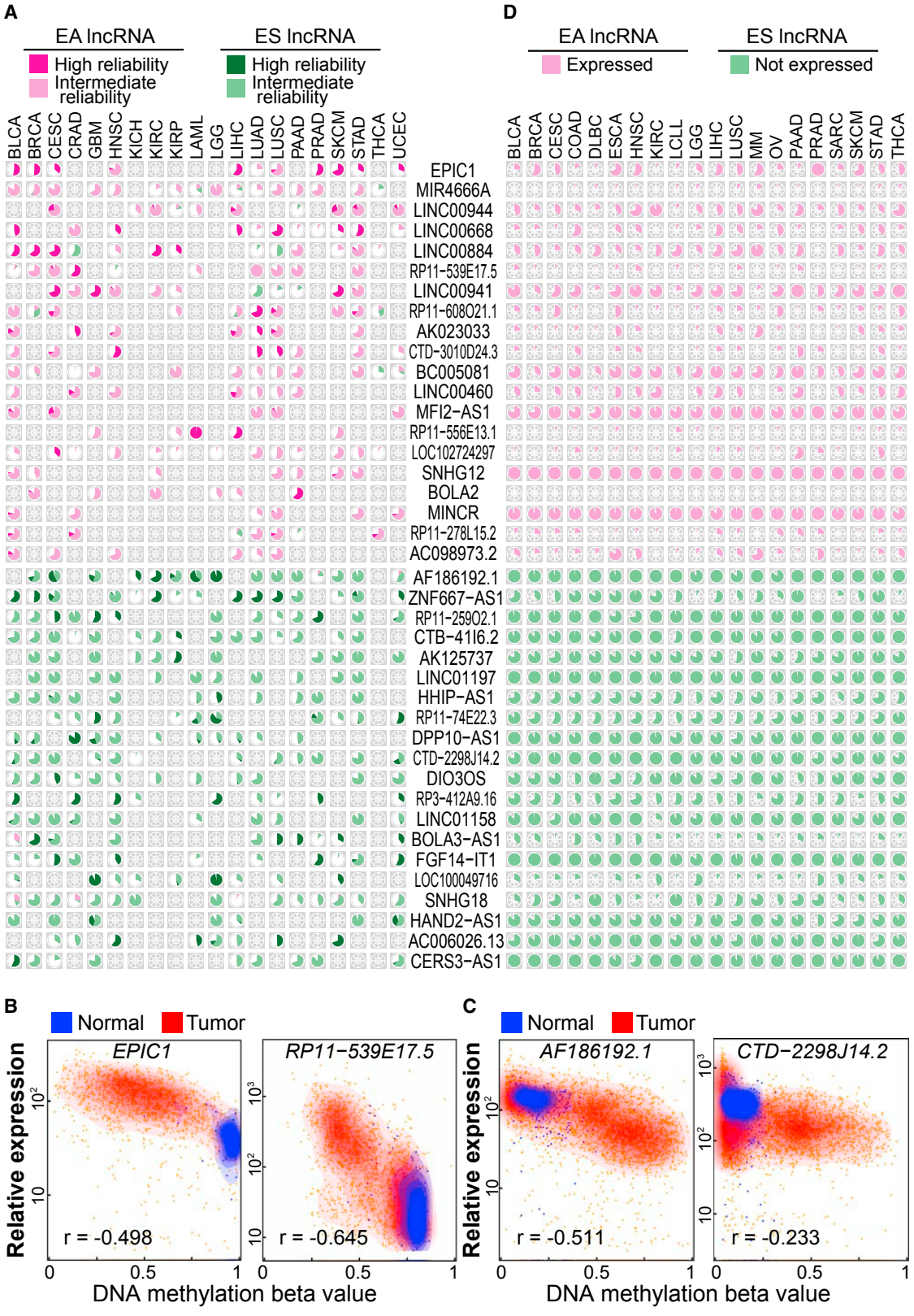
Here, we repurposed and integrated multi-dimensional genomic and epigenetic data from TCGA, Cancer Cell Line Encyclopedia (CCLE) (Barretina et al., 2012), and Catalog of Somatic Mutations in Cancer (COSMIC) (Iorio et al., 2016) projects to characterize the DNA methylation landscape of lncRNA genes across 33 cancer types. We aimed to build a detailed knowledge base and data analysis pipeline to explore DNA methylation alterations of lncRNA promoters in cancer. We hypothesize that, if some lncRNA genes are recurrently targeted by DNA methylation alterations in tumors, they may play an important role in tumor initiation and progression. By further integrating with the TCGA clinical data and somatic alterations of well-documented cancer genes, we targeted to identify and mechanistically validate lncRNAs that may have a tumor-promoting or tumor-suppressing function.

RESULTS

LncRNA Promoters Exhibit a Distinct Pattern of Epigenetic Alterations in Cancer Compared with PCG Promoters

To interrogate lncRNA DNA methylation in cancer, we developed a computational pipeline to repurpose HM450 probes to lncRNA promoters (Figures S1A and S1B). This analysis resulted in a set of 225,868 probes annotated to 28,366 genes. Specifically, 66,832 HM450 probes were annotated to 9,606 lncRNA genes (29,117 CpG islands), comprising approximately 60.4% of all lncRNAs in ENCODE annotation (Table S1). The lncRNAs that had at least one HM450 probe covering their promoters included 3,964 intergenic and 4,053 antisense lncRNA genes (Table S1). The median distances between lncRNA promoters and their nearest HM450 probes is 1,267 bp. The identified DNA methylation probes are mainly located within 3 kb regions of H3K4me3 and H3K27ac peaks of their mapped genes (Figure 1A) (ENCODE Project Consortium, 2012), suggesting that the probes indeed represent the promoter methylation status of lncRNAs and PCGs (Shlyueva et al., 2014).

We first sought to determine the lncRNA DNA methylation pattern in cancer by comparing the DNA methylation profile of lncRNA promoters between tumors and normal tissues using the TCGA Pan-Cancer database (syn4382671, Table S1). Because the CpG island hypermethylation phenotype (CIMP) has been established as one of the hallmarks in many cancer types (Baylin et al., 1986), we originally expected to identify hypermethylated tumor-suppressing lncRNAs. Intriguingly, we observed both hypermethylated and hypomethylated lncRNA promoters in breast cancer tissues (Figures 1A and S1C). This observation is in stark contrast to the PCG promoters, which were predominantly hypermethylated in breast cancer (Figure 1A). Of the intergenic lncRNAs that do not share promoters with PCGs, there were 504 intergenic lncRNA promoters showing significant hypomethylation and 639 intergenic lncRNA promoters showing significant hypermethylation in breast cancer (false discovery rate [FDR] < 0.05 and effect size > 0.2). The hypomethylation pattern of lncRNA promoters was consistently observed in another nine cancer types that also had matched normal tissues available (Figures 1B and S1D). To determine if this observation was an artifact due to bias of the HM450 microarray design, we



(legend on next page)

randomly permuted the labels of lncRNAs and PCGs for 10,000 times and generated an empirical distribution to estimate the FDR for each promoter. This analysis revealed that the lncRNA promoters were significantly hypomethylated in all ten cancer types ($p < 10^{-15}$, Kolmogorov-Smirnov test, Figure 1C).

Integrative Analysis Identified 2,123 Recurrent Epigenetically Regulated lncRNAs in 20 Cancer Types

To determine whether lncRNAs' expression is regulated by the DNA methylation changes at their promoters (e.g., hypomethylation causes overexpression), we integrated the lncRNA expression data from MiTranscriptome, which summarized the expression of 12,382 cancer-associated lncRNA transcripts using an *ab initio* assembly method in 6,475 RNA sequencing (RNA-seq) profiles, including 5,602 TCGA samples (Iyer et al., 2015). Our analysis focused on TCGA samples across 20 cancer types that have both DNA methylation and lncRNA expression data. We applied a heuristic strategy to identify the lncRNAs that are epigenetically activated (EA) or epigenetically silenced (ES) in tumors in comparison to their DNA methylation status in normal tissues. This method prioritized the lncRNAs that not only exhibited a significant difference in DNA methylation between tumors and normal tissues, but also exhibited expression changes highly correlated with their DNA methylation alterations (see details in the STAR Methods). A patient-centric matrix with DNA methylation status of 2,123 lncRNA genes across 20 cancer types was characterized, including 1,006 EA and 1,117 ES lncRNAs that showed epigenetic alteration in at least one cancer type (Table S2). The top 20 most frequently EA and ES lncRNAs are shown in Figure 2A. All the epigenetically regulated lncRNAs, with either hypomethylation or hypermethylation in tumors, exhibited a significant negative correlation (FDR < 0.01) between their expression and promoter DNA methylation status (Figures 2B and 2C). Notably, a group of the EA lncRNAs in tumors was not expressed in normal tissues (Figure S2A). This "on or off" expression pattern of EA lncRNAs potentiated them as promising diagnostic biomarkers. To further validate the methylation status of the lncRNAs and their expression in cancer, we investigated the RNA-seq and HM450 DNA methylation profiles of 455 cancer cell lines from the CCLE and COSMIC databases (Barretina et al., 2012). Among the top 40 lncRNAs, 34 (14 EA and 20 ES lncRNAs) exhibited a similar expression pattern in cancer cell lines and significantly negative correlation between their expression and promoter methylation (Figures 2D and S2B; Table S2).

Epigenetically Regulated lncRNAs Are Associated with Tumor Survival and Protein-Coding Cancer Gene Alterations

We next analyzed the association of lncRNA epigenetic status with patient survival in 20 cancer types. Twelve of the top 20 EA lncRNAs were significantly correlated with poor survival in at least 1 cancer type, while 10 of the top 20 ES lncRNAs were significantly correlated with favorable survival (Figures S2C–S2E). Among these survival-related lncRNAs are *SNHG12* and *MINCR*, which are epigenetically activated in multiple cancer types, including breast, bladder, endometrial, colorectal, and lung cancer (Figure 2A; Table S2). These lncRNAs have been documented to be overexpressed in a variety of cancer types and to play oncogenic roles in regulating cell proliferation and migration (Doose et al., 2015; Li et al., 2013; Ruan et al., 2016). To explore the relationship between lncRNA epigenetic alterations and the somatic alterations of known tumor genes, we integrated the lncRNA epigenetic alterations with the mutation and copy-number alterations of known protein-coding cancer genes in the same tumors (Vogelstein et al., 2013). Notably, the epigenetically regulated lncRNAs show a strong co-occurrence with a group of cancer gene mutations and copy-number alterations (Figure S2F; Table S2). For example, EA lncRNAs are significantly enriched in *TP53* mutated tumors in multiple cancer types (Figures S2F and S2G). By contrast, ES lncRNAs exhibit significant mutual exclusivity with *EGFR* amplifications and mutations (Figure S2F).

EPIC1 Is Epigenetically Activated and Correlated with Poor Survival in Breast Cancer

The lncRNA that is most frequently epigenetically activated in multiple cancer types is ENSG00000224271 (epigenetically induced lncRNA1 [*EPIC1*]) (Figure 2A). It is an intergenic lncRNA (CPAT coding probability = 0.004) located on chr22:q13.31. There are CpG islands within 164 bp downstream of this gene's transcription start site (Figure 3A). This lncRNA is epigenetically activated in up to 90% of tumor samples across ten cancer types, including breast cancer (Figures 2A and 2D; Table S2). Our algorithm identified three probes in HM450 mapping to the *EPIC1* CpG islands (Figure 3A). Based on the beta values of three probes, three subgroups of breast cancer were identified by the hierarchical clustering analysis in 534 breast tumors (Figure 3B). The hypermethylated subgroup includes 196 (36.7%) breast tumors and exhibits a high *EPIC1* methylation level similar to that in normal breast tissues (Figure 3B). Breast tumors of this subgroup are characterized by reduced *EPIC1* expression (Figures 3C and 3D) and an improved overall survival in comparison to the other two groups (Figure 3E). In contrast, patients whose tumors exhibit *EPIC1* hypomethylation and increased *EPIC1* expression have the worst survival (Figures 3C–3E). To determine if *EPIC1* expression is robustly associated with poor patient survival in breast cancer, we re-annotated the probes from five Affymetrix microarrays to lncRNAs and identified one probe (1563009_at) in an Affymetrix HG-U133plus2 microarray that specifically detected *EPIC1* expression. As shown in Figure 3F,

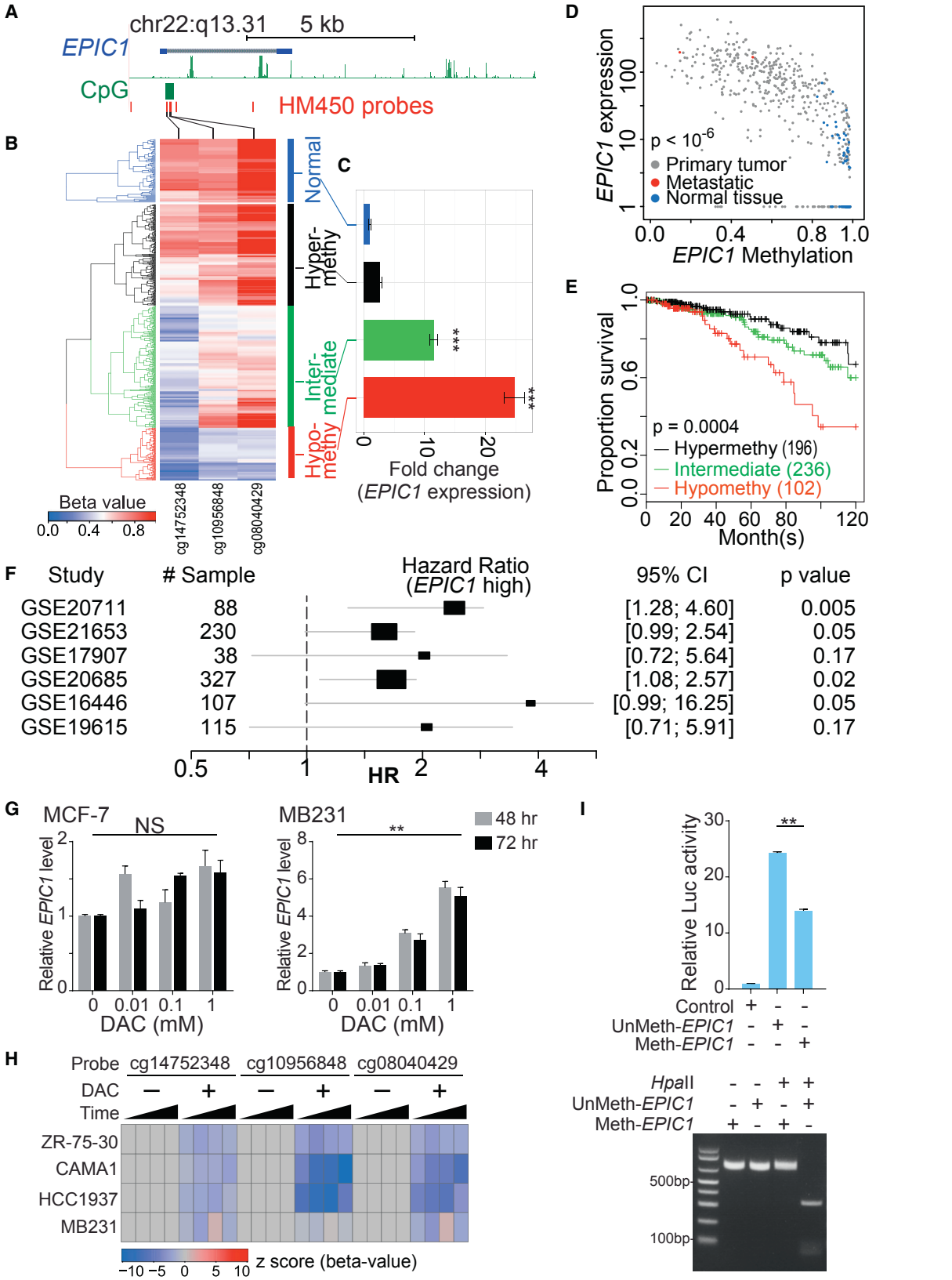
Figure 2. Epigenetic Landscape of lncRNAs in Cancer

(A) Percentages of significant EA (top panel) or ES (bottom panel) lncRNAs in 20 cancer types. Each pie chart indicates the percentage of each lncRNA epigenetic alteration in each cancer type. Purple indicates EA lncRNAs; green indicates ES lncRNAs.

(B and C) Correlation of representative EA (B) or ES (C) lncRNAs' expression and their DNA methylation level in cancer tissues (red) and normal tissues (blue). The y axis shows expression level based on RNA-seq; the x axis, DNA methylation beta values based on Infinium HM450 BeadChip.

(D) Expression of the top 20 EA (top panel) and ES (bottom panel) lncRNAs in cancer cell lines from the CCLE database. Each pie chart indicates the percentage of cell lines with the lncRNA expressed (purple, absolute read count > 0) or not expressed (green, absolute read count = 0) in each cancer type.

See also Figure S2 and Table S2.



(legend on next page)

increased expression of *EPIC1* was consistently associated with poor survival in 6 independent patient cohorts, including 905 breast tumors (Figures 3F and S3A).

Further analysis revealed that *EPIC1* epigenetic activation is significantly associated with luminal B and HER2 subtypes of breast cancer ($p < 0.001$, Figures S3B and S3C). In 119 TCGA luminal B tumors, patients with *EPIC1* epigenetic activation demonstrated significant poor survival ($p = 0.002$, Figure S3D). The association between *EPIC1* and breast cancer poor survival remains significant after adjusting cancer subtypes along with other prognostic factors including age and clinical stage (multivariate Cox regression model $p = 0.02$). In all 20 cancer types assessed, *EPIC1* epigenetic activation is also significantly correlated with poor survival in endometrial cancer patients (UCEC, Figure S2C).

Using RNA-seq and HM450 DNA methylation data in the CCLE database, we observed a significant negative correlation ($p < 0.05$) between endogenous *EPIC1* expression levels and its promoter methylation in 24 breast cancer cell lines (Figures S3E and S3F). Among them, 18 cell lines showed epigenetic activation of *EPIC1*, while 4 (i.e., MB231, HCC1937, CAMA1, and ZR-75-30) exhibited promoter hypermethylation and had low *EPIC1* expression (Figures S3E and S3F). Decitabine treatment caused a dosage- and time-dependent *EPIC1* expression and demethylation in *EPIC1* hypermethylated cell lines (e.g., MB231), but not in cells that already exhibit *EPIC1* hypomethylation and overexpression (e.g., MCF-7) (Figures 3G, 3H, and S3G). Using a similar strategy, we selected seven other EA lncRNAs based on their novelty and demonstrated that decitabine treatment significantly induced EA lncRNAs expression by decreasing the DNA methylation level of their CpG islands (Figure S3H; Table S2).

To determine if *EPIC1* is directly regulated by DNA methylation, we cloned *EPIC1*'s promoter region (including the CpG islands) and performed *in vitro* DNA methylation assay (Figure 3I). Luciferase reporter assays revealed that the unmethylated *EPIC1* promoter (unMeth-*EPIC1*) led to a significantly higher reporter activity compared with the methylated version (Meth-*EPIC1*) ($p < 0.01$, Figure 3I). Collectively, these results demonstrated that *EPIC1* is directly regulated by DNA methylation at the CpG islands in its promoter region.

EPIC1 Functions as a Potential Oncogenic lncRNA by Promoting Cell-Cycle Progression

To evaluate the oncogenic role of *EPIC1* in cancer, we analyzed the *EPIC1* expression status in 28 cell lines across 8 cancer types using qRT-PCR. In agreement with *EPIC1*'s activation in the luminal B breast cancer subtype, *EPIC1* is overexpressed in luminal breast cancer cell lines (e.g., BT-474, MB361, MCF-7, ZR-75-1, and T-47D) (Lehmann et al., 2011), along with ovarian cancer (A2780cis and OVCAR-4), pancreatic cancer (BxPC-3 and PANC-1), prostate cancer (PC-3), and leukemia (K562) cell lines (Figures S4A–S4C). We further performed 5′-RACE and 3′-RACE cloning using total RNA from MCF-7 and T-47D cells to identify functional *EPIC1* isoforms. Three splice variants of *EPIC1* were cloned, including isoform v1 (567 nt), isoform v2 (844 nt), and isoform v3 (882 nt) (Figures S4D–S4F). All of them share same exon 1 and exon 2. We designed six siRNAs targeting shared sequence of all isoforms and screened three siRNAs that can readily knockdown *EPIC1* expression (Figure S4G). *EPIC1* knockdown resulted in a decrease of cell proliferation in a time-dependent manner in luminal breast cancer cells MCF-7 and ZR-75-1 (Figures 4A–4F). Soft agar assays further demonstrated that *EPIC1* knockdown significantly inhibits the anchorage-independent growth of cancer cells (Figure 4G). Moreover, cell-cycle analysis revealed that silencing of *EPIC1* resulted in G_0/G_1 arrest in MCF-7 and ZR-75-1 cells (Figures 4C, 4F, and S4H). Next, we established stable *EPIC1* knockdown cells using lentiviral shRNAs. Both sh*EPIC1* stable cells exhibited significantly reduced cell proliferation (Figures S4I and S4J), anchorage-independent growth (Figure S4K), and *in vivo* xenograft growth (Figures 4H and 4I), compared with the shCtrl cells. These results not only suggest oncogenic activity of *EPIC1 in vivo,* but also provide a potential therapeutic target for breast cancer treatment.

EPIC1 Is a Nuclear lncRNA that Regulates MYC Targets

Cell fractionation PCR and subcellular RNA-seq analyses revealed that *EPIC1* RNA is predominately located in the nucleus (Figures 5A, S5A, and S5B), suggesting that *EPIC1* might play a role in transcriptional regulation and chromatin interactions (Batista and Chang, 2013). To explore this possibility, RNA-seq analyses were performed on MCF-7 cells transfected with two

Figure 3. Expression Level of *EPIC1* Is Regulated by DNA Methylation and Associated with Poor Survival in Breast Cancer Patients

(A) The locations of *EPIC1* gene (blue), CpG islands (green), and HM450 probes (red) in GRCh37 reference human genome (chr22:48,027,423-48,251,349).

(B) Heatmap with beta values of DNA methylation obtained from three *EPIC1* HM450 probes in breast normal tissues and tumors. Three subgroups were identified using a hierarchical clustering analysis in tumors. Black, hypermethylation; green, intermediate; red, hypomethylation. *EPIC1*'s DNA methylation in normal tissues (blue) is shown as control. Full IDs of *EPIC1* HM450 probes are cg10956848, cg14752348, and cg08040429.

(C) Relative *EPIC1* expression in three subgroups above, compared with the level in normal tissues, respectively. The error bars represent standard errors. ***$p < 0.001$.

(D) Correlation of *EPIC1* expression with *EPIC1* DNA methylation status in breast cancer and normal tissues. Probe cg08040429 represents the DNA methylation status.

(E) Kaplan-Meier survival curve represents the proportion survival of breast cancer patients with three subgroups above.

(F) Forest plot of *EPIC1*'s association with survival in six independent breast cancer cohorts. *EPIC1*'s expression is measured by Affymetrix 1563009_at (HG-U133_Plus_2).

(G) qRT-PCR analysis of *EPIC1* expression in MCF-7 and MB231 cells treated with decitabine (DAC).

(H) *EPIC1* methylation status detected by the same three probes (B) in breast cancer cell lines treated with decitabine. Differences in z score-transformed beta value are shown.

(I) Reporter assay of methylated and unmethylated *EPIC1* promoters (top). *In vitro* DNA methylation status of *EPIC1* promoters was confirmed by HpaII restriction enzyme (bottom).

Error bars indicate mean ± SD, n = 3 for technical replicates. **$p < 0.01$. NS, not significant.

See also Figure S3.

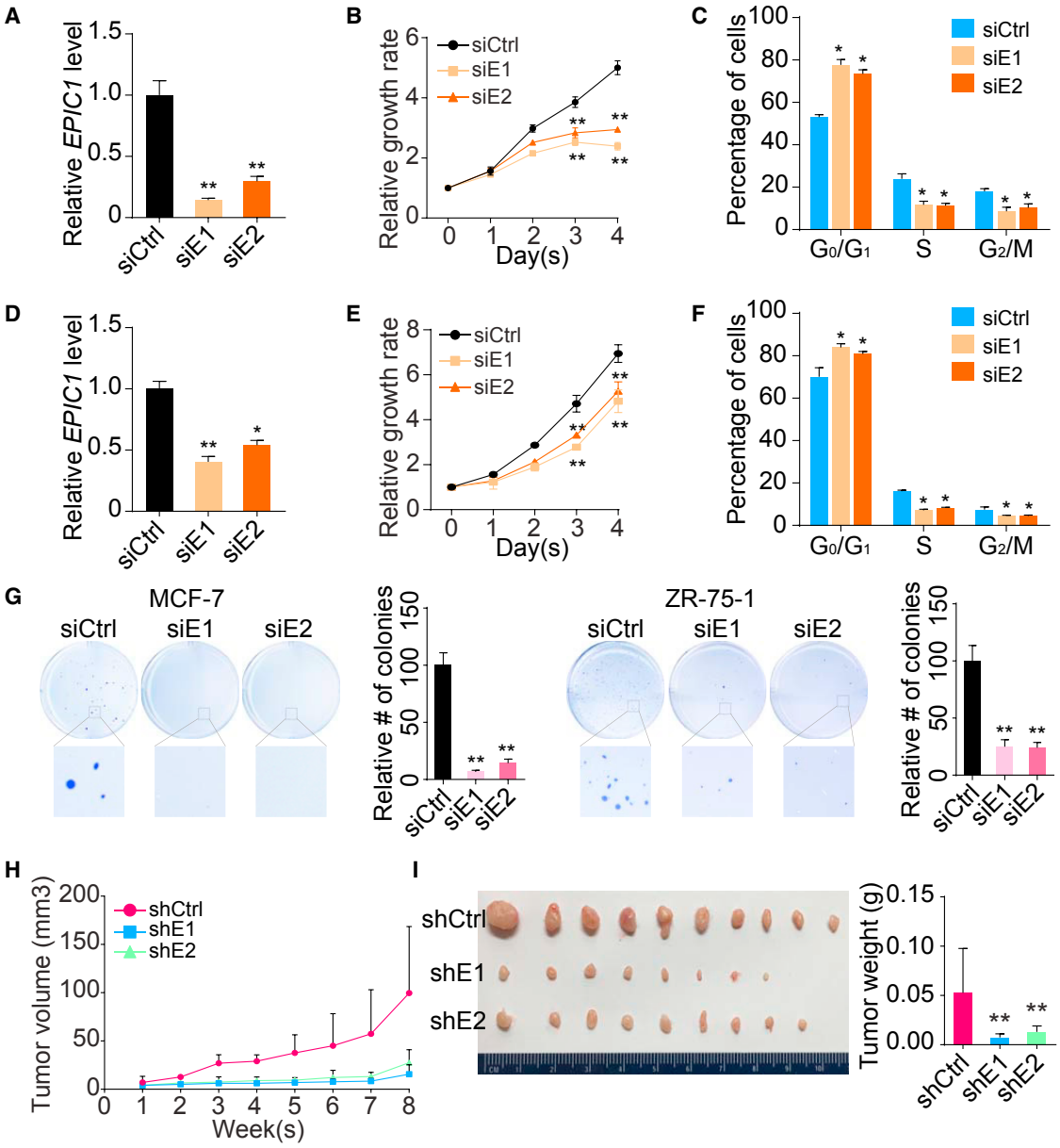


Figure 4. *EPIC1* Functions as an Oncogenic lncRNA in Breast Cancer

(A–C) qRT-PCR analysis of *EPIC1* (A), MTT assay (B), and cell-cycle analysis (C) in MCF-7 cells treated with *EPIC1* siRNAs (siE1 and siE2).

(D–F) qRT-PCR analysis of *EPIC1* (D), MTT assay (E), and cell-cycle analysis (F) in ZR-75-1 cells treated with *EPIC1* siRNAs.

(G) Anchorage-independent colony formation assays of MCF-7 (left) and ZR-75-1 (right) cells treated with *EPIC1* siRNAs.

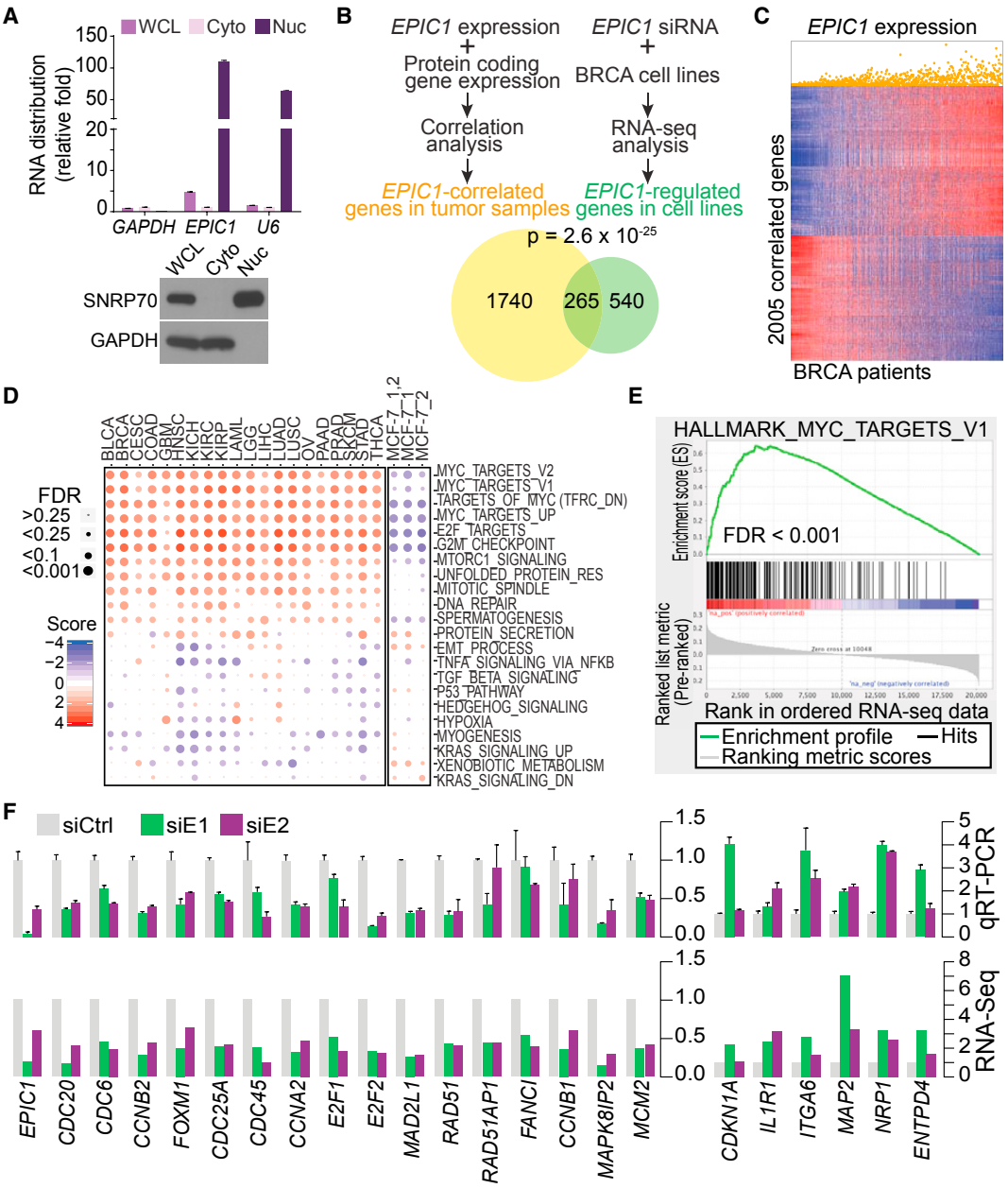
(H) Quantification of tumor growth in xenograft mouse models bearing with stable *EPIC1* knockdown (shE1 and shE2) or control (shCtrl) MCF-7 cells.

Error bars indicate means ± SD, n = 3 for technical replicates. *$p < 0.05$, **$p < 0.01$.

(I) Representative tumor size (left), and quantification of tumor weight (right) from xenograft mouse models. Data are presented as means ± SD (n = 10). **$p < 0.01$.

See also Figure S4.

siRNAs targeting *EPIC1* individually or pooled. We have confirmed that both siRNAs can readily knockdown the level of nuclear *EPIC1* RNA (Figure S5C). To exclude possible off-target effects on gene expression associated with single siRNAs, we focused only on genes regulated in the same direction in all three transfection experiments. *EPIC1* knockdown in MCF-7 cells resulted in the regulation of 805 genes (upregulation of 317 genes and downregulation of 488 genes) (Figure 5B; Table S3), which are highly overlapped with 2,005 *EPIC1*-associated genes that were significantly correlated with *EPIC1* expression across 559 TCGA breast tumors ($p = 2.6 \times 10^{-25}$, Figures 5B and 5C). This overlap was even higher in the pathway analysis. Gene set enrichment analysis (GSEA) analysis showed that cell-cycle-related biological processes such as "MYC targets," "G2M checkpoint," and "E2F targets" were significantly enriched in the *EPIC1*-associated genes in 17 out of 20 cancer types



(legend on next page)

(Figure 5D; Table S3). The same cellular processes were enriched in the *EPIC1*-regulated genes in MCF-7 cells (Figure 5D; Table S3). Among them, the MYC pathway/targets are prominent gene sets enriched with *EPIC1*-regulated genes in both tumor samples and cell lines (Figure 5E). For example, the MYC targets *CDC45*, *CDC20*, and *CCNA2* were significantly downregulated by *EPIC1* knockdown. Moreover, *CDKN1A* (encoding the p21 protein) was significantly induced after *EPIC1* knockdown (Figures 5F, 5G, and S5D). p21 is a well-established negative regulator of cell-cycle progression at G_1 and S phase that is directly inhibited by MYC (Gartel and Radhakrishnan, 2005). These observations are consistent with our observation that *EPIC1* knockdown resulted in cancer cells' arrest at G_0/G_1 phase. Similarly, in MCF-7 and ZR-75-1 cells, MYC knockdown also led to a pattern of MYC target expression and cell growth comparable with *EPIC1* knockdown (Figures 5G and S5E–S5H). This suggested that the oncogenic role of *EPIC1* may be associated with MYC protein.

EPIC1 Interacts with the 148–220 Amino Acid Region of MYC through Its 129–283 nt Sequence

To study the interaction between *EPIC1* RNA and MYC protein, we overexpressed each of three *EPIC1* isoforms (i.e., v1, v2, and v3) with Flag-tagged MYC protein in 293T cells, and performed RNA immunoprecipitation (RIP) assay. This analysis revealed that *EPIC1* isoforms v1 and v2 could be enriched by MYC RIP (Figure S6A). In v1 or v2 isoforms overexpressing MCF-7 cells, only the v1 isoform could regulate MYC target genes (Figures S6B and S6C). We further observed that overexpression of the *EPIC1* v1 isoform promoted G_1 phase progression and *in vivo* xenograft growth (Figures S6D–S6F). It is apparent to us that the v1 isoform is the functional isoform of *EPIC1* gene in breast cancer. We therefore used isoform v1 (567 nt) as the reference sequence of *EPIC1* in the following study.

RNA pull-down assay showed that MYC protein could be co-precipitated by an *in-vitro*-transcribed biotinylated *EPIC1* sense transcript, but not by the *EPIC1* antisense transcript (Figure 6A). MYC RIP with cell lysates from MCF-7 cells was then performed to confirm the interaction between endogenous *EPIC1* and MYC protein (Figures 6B, S6G, and S6H). A well-documented MYC interacting lncRNA, *PVT1* (Tseng et al., 2014), was included as positive control and could also be enriched by MYC RIP (Figure 6B). Further *in vitro* binding assay using *in-vitro*-transcribed *EPIC1* RNA and recombinant His-tagged MYC protein demonstrated that *EPIC1* binds directly to MYC protein (Figure 6C). To map the *EPIC1* functional motifs corresponding to MYC binding, we conducted an *in vitro* RNA pull-down assay using a series of truncated *EPIC1* fragments. This analysis revealed that nucleotides 1–358 of *EPIC1* (*EPIC1* 1–358 nt) are sufficient to interact with MYC protein, while other *EPIC1* truncated fragments could not (Figure 6D). To map with greater precision the sequence of *EPIC1* that binds to MYC, we further designed seven truncated or deletion mutants of the *EPIC1* 1–358 nt region and revealed that three deletion mutants (Δ121-180 nt, Δ181-240 nt, and Δ241-300 nt) can abolish *EPIC1* binding to MYC protein. Deletion of all three regions (129–283 nt) also abolished *EPIC1*'s interaction with MYC protein (named as *ΔMYC-EPIC1*; Figures 6E and 6F). These data suggested that the *EPIC1* 129–283 nt region is necessary for *EPIC1*'s binding to the MYC protein. MYC protein domain mapping studies revealed that *EPIC1* binds the 148–220 amino acid (aa) region of MYC, which is not overlapped with the well-characterized transcriptional activation domain and basic-helix-loop-helix domain of MYC protein (Luscher, 2001; von der Lehr et al., 2003) (Figures 6G and 6H). Deletion of the 148–220 aa region of MYC protein (named as ΔEPIC1-MYC) abolished its interaction with *EPIC1* (Figures 6G and 6H). Collectively, our findings demonstrated that *EPIC1* interacts with the 148–220 aa region of MYC through its 129–283 nt sequence.

The Oncogenic Role of *EPIC1* Partially Depends on Its Regulation of MYC Occupancy on Target Promoters

With the observation that *EPIC1* directly interacts with MYC, we further analyzed the effect of *EPIC1* on MYC target gene reporters (e.g., *p21* and *CCNA2* promoters) in MCF-7 cells. The reporter assays revealed that knockdown of either *EPIC1* or *MYC* significantly regulates p21-Luc and CCNA2-Luc reporter luciferase activities (Figure 7A). These observations indicate that *EPIC1* directly regulates the expression of MYC targets through their promoter regions. Interestingly, *EPIC1* knockdown had little effect on the expression of MYC (Figure 5G), which led to our hypothesis that *EPIC1* may regulate the transcriptional activity of the MYC protein.

To test this hypothesis, we performed an integrated analysis on MYC chromatin immunoprecipitation sequencing (ChIP-seq) data (Lee et al., 2012) and RNA-seq data of *EPIC1* knockdown MCF-7 cells. Among 805 *EPIC1*-regulated genes, 785 have robust MYC occupancy on their promoters in two biological replicates of MCF-7 ChIP-seq data. Interestingly, we did not observe a significant correlation between global MYC binding

Figure 5. *EPIC1* Is a Nuclear lncRNA Regulating MYC Targets Expression

(A) qRT-PCR analysis of *EPIC1* expression (top) and western blot (bottom) of subcellular fractionation in MCF-7cells. *GAPDH* and *U6* RNA served as markers for cytoplasmic and nuclear gene localization, respectively. SNRP70 and GAPDH served as a specific nuclear and cytoplasmic marker to whole-cell lysates (WCL), cytoplasmic (Cyto), and nuclear fractionation (Nuc). Error bars indicate mean ± SD, n = 3 for technical replicates.

(B) Schematic of the identification of *EPIC1* correlated genes in breast tumors from TCGA (yellow), and genes potentially regulated by *EPIC1* in MCF-7 cells (green).

(C) Co-expression analysis showing that *EPIC1* expression is associated with 2,005 genes in 559 patients with breast cancer (BRCA). Each column represents one patient.

(D) Gene set enrichment analysis (GSEA) of the *EPIC1*-related pathways in 20 cancer types (left panel) and *EPIC1* knockdown MCF-7 cells (right panel). The heatmap indicates the GSEA scores.

(E) Association between the enrichment of MYC targets and *EPIC1* expression in breast tumors by GSEA analysis (D).

(F) *EPIC1*-regulated gene expression by qRT-PCR analysis (top) and RNA-seq (bottom). Error bars indicate mean ± SD, n = 3 for technical replicates.

(G) Western blot of MYC-regulated targets in MCF-7 (left) and ZR-75-1 (right) cells treated with *EPIC1* and MYC siRNAs.

See also Figure S5 and Table S3.

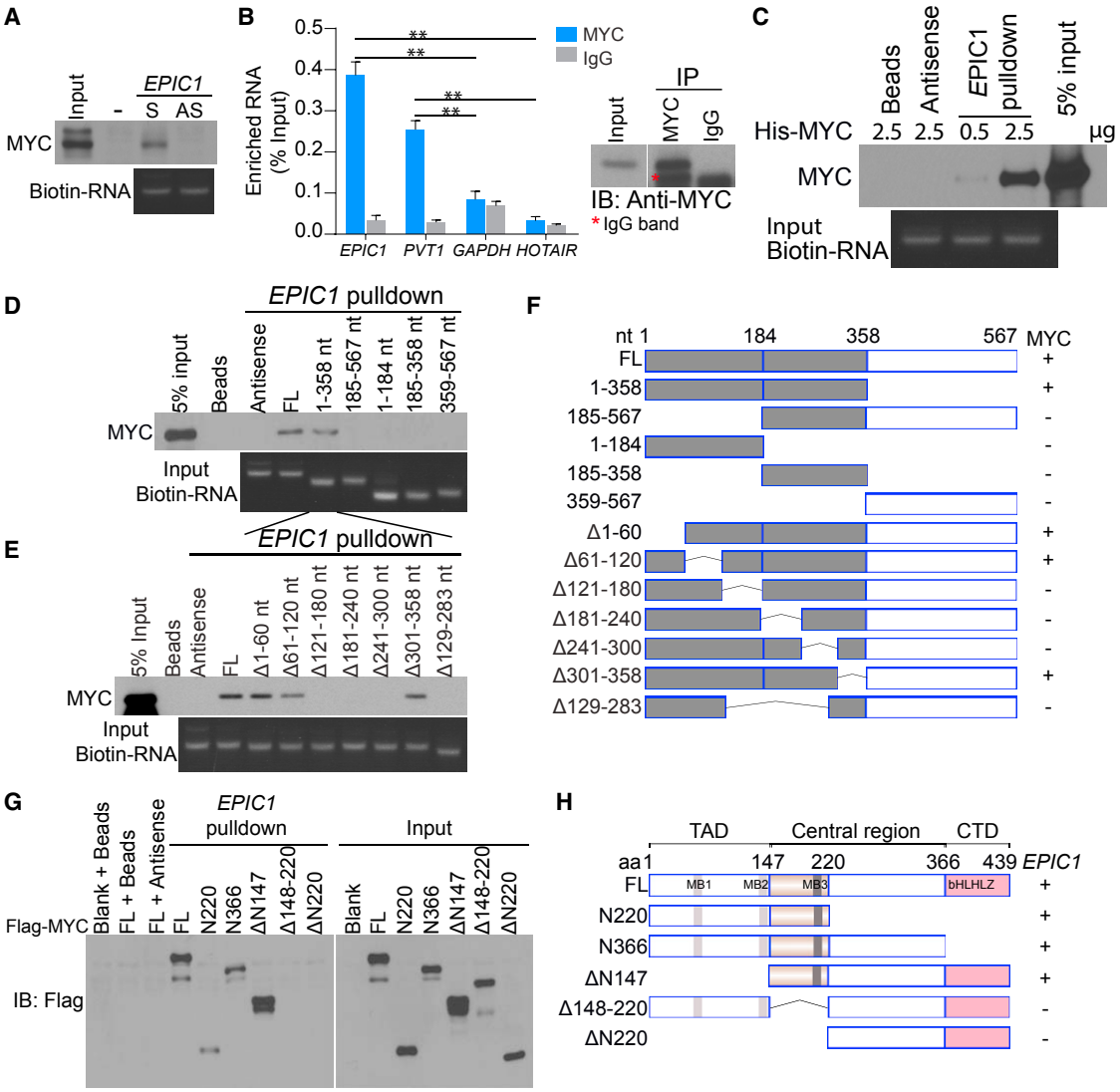


Figure 6. *EPIC1* Binds Directly with MYC

(A) Western blot of MYC proteins retrieved by *in-vitro*-transcribed biotinylated *EPIC1* from MCF-7 cell nuclear extracts. Antisense *EPIC1* was used as a negative control. S, sense strand; AS, antisense strand.

(B) qRT-PCR analysis of *EPIC1* and *PVT1* enriched by MYC proteins in MCF-7 cells. Western blot of MYC is shown (right). *HOTAIR* and *GAPDH* served as negative controls. Error bars indicate mean ± SD, n = 3 for technical replicates. **p < 0.01.

(C) Western blot of recombinant MYC proteins retrieved by *EPIC1* RNA in *in vitro* binding assay. *EPIC1* antisense was used as a negative control.

(D) Western blot of MYC pulled down by truncated *EPIC1*.

(E) Mapping of the MYC binding region within the 1–358 region of *EPIC1*.

(F) Schematic of truncated or deletion mutants of *EPIC1*. The MYC binding capability is shown (right).

(G) Western blot of truncated MYC proteins retrieved by *in-vitro*-transcribed *EPIC1*.

(H) Schematic of truncated MYC protein. The *EPIC1* binding capability is shown. TAD, N-terminal transactivation domain; MB1-3, MYC boxes 1–3; bHLHLZ, basic-helix-loop-helix-leucine zipper domain; CTD, C-terminal domain.

See also Figure S6.

affinity and differential expression (i.e., fold change) after *EPIC1* knockdown in MCF-7 cells, suggesting that *EPIC1* may regulate MYC's occupancy on a specific group of targets. By further considering previously validated MYC targets (Li et al., 2003; Zeller et al., 2006), we identified 40 possible targets of the *EPIC1*-MYC regulatory axis (Figures 7B and S7A; Table S4). ChIP-qPCR were performed and validated that *EPIC1* knockdown significantly reduces MYC's occupancies on the promoters of 26 targets, including *CDKN1A* (*p21*), *CCNA2, CDC20,* and *CDC45* (Figures 7C, S7B, and S7C). It is known that MYC binds to DNA and functions as a transcription factor by heterodimerization with another transcription factor, MAX (Amati et al., 1993; Blackwood and Eisenman, 1991). MYC and MAX Co-IP assay in MCF-7 cells revealed that *EPIC1* knockdown could moderately reduce the formation of MYC-MAX complexes (Figure S7D). Moreover, overexpression of *EPIC1*, but not *ΔMYC-EPIC1*, could

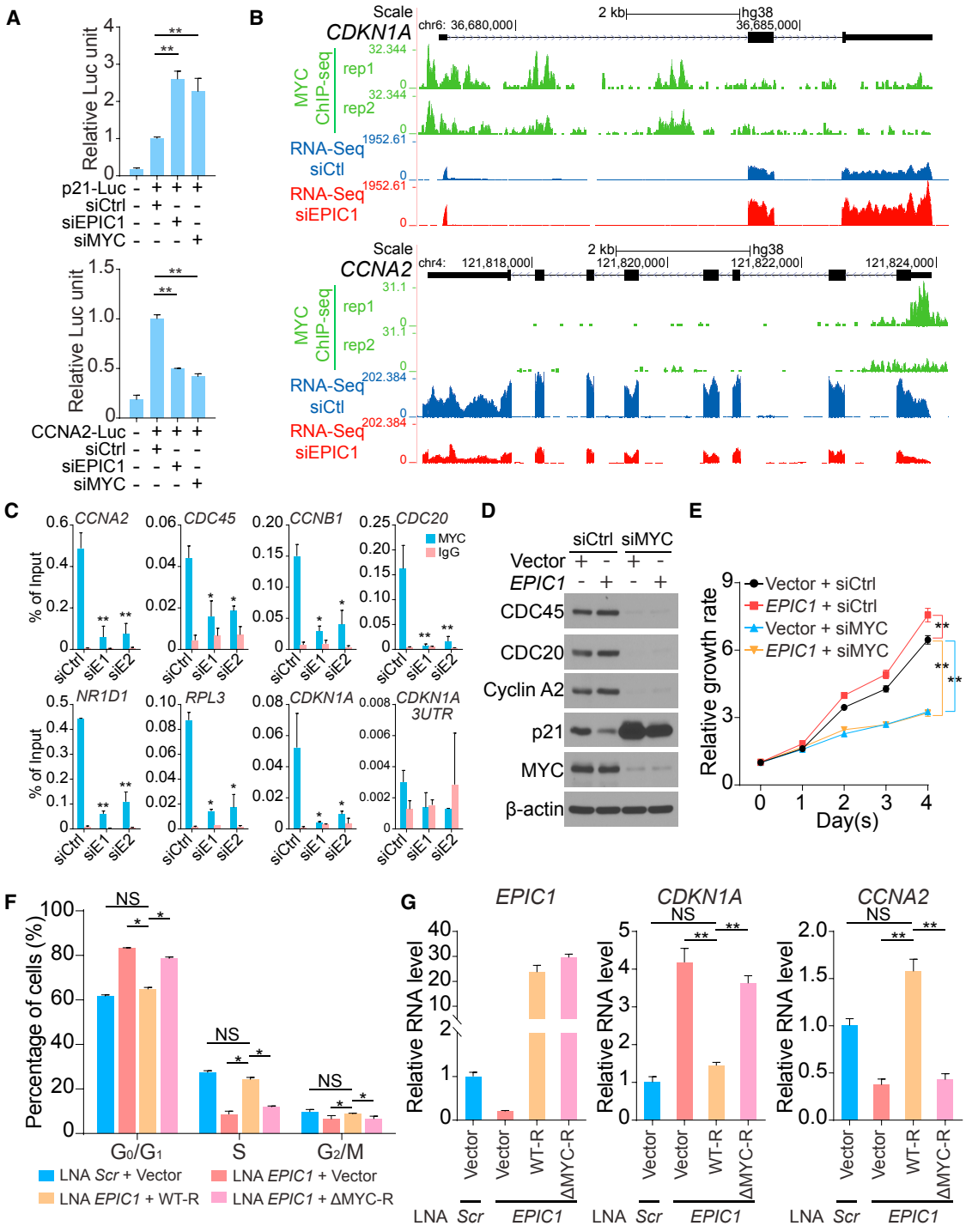


Figure 7. MYC Is Required for the Regulatory Role of *EPIC1* in Cancer

(A) Reporter assay of *CDKN1A* (p21) and *CCNA2* (Cyclin A2) promoters.

(B) Alignment of two biological replicates of MYC ChIP-seq in MCF-7 cells (green) and RNA-seq from siCtrl (blue) and siEPIC1 (red) RNA-treated MCF-7 cells. *CDKN1A* and *CCNA2* genomic locus are shown.

(C) ChIP-qPCR analysis of MYC occupancy on the promoters of target genes in MCF-7 cells treated with *EPIC1* siRNAs.

(D and E) Western blot of MYC targets (D) and MTT assay (E) after treatment with *MYC* siRNAs in MCF-7 cells with stable overexpression of *EPIC1* and empty vector.

(F and G) Cell-cycle analysis (F) and qRT-PCR analysis of *EPIC1*, *CDKN1A*, and *CCNA2* level (G) in MCF-7 cells transfected with LNA against *EPIC1* followed by overexpression of indicated vectors.

Error bars indicate mean ± SD, n = 3 for technical replicates. *p < 0.05, **p < 0.01. NS, not significant.

See also Figure S7 and Table S4.

enhance the reporter luciferase activities mediated by MYC and MAX (Figure S7E). These results suggest that *EPIC1* promotes MYC's occupancy on *EPIC1*-regulated genes through its 129–283 nt sequence (i.e., MYC-binding sequence).

To further determine the role of the *EPIC1*-MYC regulatory axis in cancer, we performed the MYC knockdown in *EPIC1* stably overexpressing MCF-7 cells, and observed that *EPIC1* regulation of cell proliferation and MYC target expression were attenuated by MYC knockdown (Figures 7D and 7E). Overexpression of MYC, but not *EPIC1*-binding-deficient mutant MYC proteins (ΔEPIC1-MYC), regulates CCNA2 and p21 expression (Figure S7F). We further depleted the endogenous *EPIC1* expression using locked nucleic acid (LNA) in MCF-7 cells, followed by overexpression of either LNA-resistant wild-type *EPIC1* (*WT-R-EPIC1*) or deletion mutant of 129–283 nt MYC binding region (*ΔMYC-R-EPIC1*). Similar to *EPIC1* siRNA treatment, LNA knockdown of *EPIC1* significantly caused G_1 arrest of MCF-7 cells, which could be rescued by reintroduction of full-length *EPIC1*, but not *ΔMYC-EPIC1* (Figures 7F and S7G). The expression of full-length and the truncated *EPIC1s* was confirmed to be comparable levels to rule out the influence of transfection efficiency (Figure 7G). Consistently, LNA knockdown of *EPIC1* also curtailed the expression of MYC target genes. Reintroduction of wild-type *EPIC1*, but not *ΔMYC-EPIC1*, was able to rescue the regulation of these genes (Figure 7G). These results suggested that the oncogenic role of *EPIC1* is at least in part dependent on its interaction with the MYC protein.

DISCUSSION

Previous studies, by repurposing copy-number and gene expression microarray data, have successfully identified the copy-number alterations (Hu et al., 2014; Leucci et al., 2016; Tseng et al., 2014; Yan et al., 2015) and expression alterations of lncRNA (Du et al., 2013) in cancer. In the present study, we repurposed and integrated multi-dimensional genomic and epigenetic data from 6,475 tumor samples and 455 cancer cell lines in the TCGA and CCLE projects. These data were remapped/realigned to 9,606 annotated human lncRNAs to comprehensively characterize the lncRNA DNA methylation landscape in cancer. Our analyses demonstrate that integrating HM450 microarray and RNA-seq data is a cost-effective strategy to research the DNA methylation regulation of lncRNA genes given the large number of HM450 and RNA-seq datasets available in public repositories. Our study has revealed that lncRNAs can be epigenetically activated in tumors by loss of DNA methylation in the promoter region, which is in stark contrast to the well-documented CIMP phenotype of PCGs in tumors. By further integrating with the protein-coding cancer gene alterations in same tumors, we observed that lncRNA epigenetic activation exhibited a strong co-occurrence with *TP53* mutation in multiple cancer types. Emerging evidence has demonstrated that p53 is a master regulator of lncRNAs' expression in cancer (Sanchez et al., 2014; Schmitt et al., 2016). Future study is warranted to determine whether loss of DNA methylation makes the promoters of these lncRNAs accessible to transcription factors, such as p53, and leads to transcriptional activation.

We hypothesize that, if some lncRNAs are recurrently targeted by epigenetic alterations in tumors, they may play an important role in tumor initiation and progression. Indeed, the epigenetically regulated lncRNAs identified in this study include a number of known cancer-related lncRNAs, such as *KCNQ1OT1* (Engel et al., 2000), *MEG3* (Zhou et al., 2012), *MINCR* (Li et al., 2013), *HOTAIR* (Gupta et al., 2010), and *WT1-AS* (Hancock et al., 2007). Consistent with their somatic DNA methylation alterations identified in this study, germline epigenetic defects in some of those lncRNAs have been documented to cause predisposition to Wilms tumor (Scott et al., 2008) and pediatric adrenocortical tumors (Wijnen et al., 2012).

Encouraged by the recapitulation of documented cancer-related lncRNAs, we mechanistically validated the most frequently EA lncRNA, *EPIC1*, as a potential oncogene. We have demonstrated that *EPIC1* interacts with MYC protein through its 129–283 nt region and increases MYC occupancy on *EPIC1*-regulated genes. The oncogenic role of MYC has been well documented in cancer initiation and progression (Dang, 2012). As an oncogene, *MYC* can be activated by multiple mechanisms in cancer. Chromosomal rearrangement is believed to be the most common genetic alteration of *MYC* (Dang, 2012). Other MYC activation mechanisms include transcriptional regulation, mRNA stabilization, and protein overexpression and stabilization (Kress et al., 2015). Emerging evidence has uncovered lncRNA's role for MYC activation in cancers. Three recent reports identified lncRNA *CCAT1-L* (colorectal cancer-associated transcript 1), *GHET1* (gastric carcinoma highly expressed transcript 1), and *PCGM1* (prostate cancer gene expression marker 1) to be involved in modulating the transcription (Xiang et al., 2014) or RNA stability (Yang et al., 2014) of MYC in colorectal, gastric, and prostate cancers (Hung et al., 2014). Another study demonstrated lncRNA *PVT1* (plasmacytoma variant translocation 1) as an oncogenic lncRNA that interacts and stabilizes the MYC protein (Tseng et al., 2014).

However, little is known about whether and how lncRNAs regulate the transcriptional activity of MYC. MYC protein alone cannot form a homodimer nor bind to DNA *in vivo*. In most cases, MYC heterodimerizes with a partner protein, MAX (Amati et al., 1993; Blackwood and Eisenman, 1991) via a basic-helix-loop-helix-leucine zipper domain. The MYC-MAX complex binds directly to DNA sequence (CACA/GTG), which is a subset of the general E-box (CACGTG) DNA recognition sequence and functions as transcriptional activator or repressor (Blackwood and Eisenman, 1991; Luscher, 2001). It has been reported that MYC is bound to ~25,000 sites in the human genome (Cawley et al., 2004; Fernandez et al., 2003). Among those *in vivo* MYC binding sites, only a small set of sites have an MYC-MAX consensus CACA/GTG sequence (Fernandez et al., 2003). One reason for this discrepancy is that MYC can be recruited to non-canonical binding sites by other transcription factors. For example, MYC can interact with Miz1, which recruits MYC to its core promoter sequences that lack an MYC-MAX binding motif (Peukert et al., 1997). Other proteins, which recruit MYC to their cognate DNA binding sites, include specificity protein-1 (Gartel et al., 2001), nuclear factor Y (Izumi et al., 2001), transcription factor II-I) (Roy et al., 1993), and yingyang-1 (Shrivastava et al., 1993). In the current study, our results suggest that *EPIC1* specifically regulates MYC's occupancy on a subset of MYC targets. Our results also showed that *EPIC1* can moderately enhance

MYC-MAX interaction. It is possible that *EPIC1* only influences MYC's occupancy on canonical MYC-MAX binding sites, but not the non-canonical MYC binding sites mediated by other "tethering factors." Another possible explanation is that *EPIC1* may function as a "guide" RNA to facilitate MYC-MAX's regulation on specific targets by directly binding to double-strand DNA. Future study is required to further define how *EPIC1* regulates MYC's occupancy on these specific MYC targets.

In summary, the establishment of a detailed knowledge base of the DNA methylation-altered lncRNAs in cancer will facilitate the identification of cancer-driving lncRNAs. Moreover, the mechanistic characterization of *EPIC1* and its functional cross-talk with the well-established oncogene *MYC* may help to pave the way to develop cancer therapies that target MYC through its interaction with *EPIC1*. The strong prognostic association of *EPIC1*, the robust tumor growth suppression by the *EPIC1* knockdown, and the illustration of *EPIC1*'s mechanism to promote breast cancer will shed light on the future development of lncRNA-based breast cancer therapies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and five tables and can be found with this article online at https://doi.org/10.1016/j.ccell.2018.03.006.

ACKNOWLEDGMENTS

This study was supported by a grant from the Shear Family Foundation (to D.Y.), a Career Development Award of RPCI-UPCI Ovarian Cancer SPORE (P50 CA159981; to D.Y.), and a grant from the Elsa U. Pardee Foundation (to D.Y.). This project is supported in part by award P30CA047904. We thank The Center for Simulation and Modeling (SaM) at the University of Pittsburgh for computing assistance. We thank The Center for Biologic Imaging (CBI) at the University of Pittsburgh for imaging assistance. We thank Drs. Nara Lee, Barry Gold, Anil Sood, Wei Zhang, and Ilya Shmulevich for internal critical review.

AUTHOR CONTRIBUTIONS

Conceptualization, D.Y., B.Y., and Z.W.; Methodology, Z.W., B.Y., M.Z., and D.Y.; Formal Analysis, B.Y., Z.W., M.Z., W.G., Z.Y.W., and D.Y.; Investigation, Z.W., B.Y., M.Z., W.G., Z.Y.W., Y.W., L.J., W.X., S.L., and D.Y.; Resources, B.Y. and M.Z.; Writing – Original Draft, Z.W., B.Y., M.Z., W.X., and D.Y.; Writing – Review & Editing, Z.W., B.Y., M.Z., W.G., Z.Y.W., W.X., S.L., and D.Y.; Supervision & Funding Acquisition, D.Y.

DECLARATION OF INTERESTS

Michael Seiler, Peter G. Smith, Ping Zhu, Silvia Buonamici, and Lihua Yu are employees of H3 Biomedicine, Inc. Parts of this work are the subject of a patent application: WO2017040526 titled "Splice variants associated with neomorphic sf3b1 mutants." Shouyoung Peng, Anant A. Agrawal, James Palacino, and Teng Teng are employees of H3 Biomedicine, Inc. Andrew D. Cherniack, Ashton C. Berger, and Galen F. Gao receive research support from Bayer Pharmaceuticals. Gordon B. Mills serves on the External Scientific Review Board of Astrazeneca. Anil Sood is on the Scientific Advisory Board for Kiyatec and is a shareholder in BioPath. Jonathan S. Serody receives funding from Merck, Inc. Kyle R. Covington is an employee of Castle Biosciences, Inc. Preethi H. Gunaratne is founder, CSO, and shareholder of NextmiRNA Therapeutics. Christina Yau is a part-time employee/consultant at NantOmics. Franz X. Schaub is an employee and shareholder of SEngine Precision Medicine, Inc. Carla Grandori is an employee, founder, and shareholder of SEngine Precision Medicine, Inc. Robert N. Eisenman is a member of the Scientific Advisory Boards and shareholder of Shenogen Pharma and Kronos Bio. Daniel J. Weisenberger is a consultant for Zymo Research Corporation. Joshua M. Stuart is the founder of Five3 Genomics and shareholder of NantOmics. Marc T. Goodman receives research support from Merck, Inc. Andrew J. Gentles is a consultant for Cibermed. Charles M. Perou is an equity stock holder, consultant, and Board of Directors member of BioClassifier and GeneCentric Diagnostics and is also listed as an inventor on patent applications on the Breast PAM50 and Lung Cancer Subtyping assays. Matthew Meyerson receives research support from Bayer Pharmaceuticals; is an equity holder in, consultant for, and Scientific Advisory Board chair for OrigiMed; and is an inventor of a patent for EGFR mutation diagnosis in lung cancer, licensed to LabCorp. Eduard Porta-Pardo is an inventor of a patent for domainXplorer. Han Liang is a shareholder and scientific advisor of Precision Scientific and Eagle Nebula. Da Yang is an inventor on a pending patent application describing the use of antisense oligonucleotides against specific lncRNA sequence as diagnostic and therapeutic tools. Yonghong Xiao was an employee and shareholder of TESARO, Inc. Bin Feng is an employee and shareholder of TESARO, Inc. Carter Van Waes received research funding for the study of IAP inhibitor ASTX660 through a Cooperative Agreement between NIDCD, NIH, and Astex Pharmaceuticals. Raunaq Malhotra is an employee and shareholder of Seven Bridges, Inc. Peter W. Laird serves on the Scientific Advisory Board for

AnchorDx. Joel Tepper is a consultant at EMD Serono. Kenneth Wang serves on the Advisory Board for Boston Scientific, Microtech, and Olympus. Andrea Califano is a founder, shareholder, and advisory board member of DarwinHealth, Inc. and a shareholder and advisory board member of Tempus, Inc. Toni K. Choueiri serves as needed on advisory boards for Bristol-Myers Squibb, Merck, and Roche. Lawrence Kwong receives research support from Array BioPharma. Sharon E. Plon is a member of the Scientific Advisory Board for Baylor Genetics Laboratory. Beth Y. Karlan serves on the Advisory Board of Invitae.

Received: June 25, 2017
Revised: February 13, 2018
Accepted: March 5, 2018
Published: April 2, 2018

REFERENCES

Amati, B., Brooks, M.W., Levy, N., Littlewood, T.D., Evan, G.I., and Land, H. (1993). Oncogenic activity of the c-Myc protein requires dimerization with Max. Cell *72*, 233–245.

Amin, V., Harris, R.A., Onuchic, V., Jackson, A.R., Charnecki, T., Paithankar, S., Lakshmi Subramanian, S., Riehle, K., Coarfa, C., and Milosavljevic, A. (2015). Epigenomic footprints across 111 reference epigenomes reveal tissue-specific epigenetic regulation of lincRNAs. Nat. Commun. *6*, 6370.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehar, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607.

Batista, P.J., and Chang, H.Y. (2013). Long noncoding RNAs: cellular address codes in development and disease. Cell *152*, 1298–1307.

Baylin, S.B., Hoppener, J.W., de Bustros, A., Steenbergh, P.H., Lips, C.J., and Nelkin, B.D. (1986). DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. Cancer Res. *46*, 2917–2922.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol *57*, 289.

Blackwood, E.M., and Eisenman, R.N. (1991). Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Science *251*, 1211–1217.

Brennan, C.W., Verhaak, R.G., McKenna, A., Campos, B., Noushmehr, H., Salama, S.R., Zheng, S., Chakravarty, D., Sanborn, J.Z., Berman, S.H., et al. (2013). The somatic genomic landscape of glioblastoma. Cell *155*, 462–477.

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J., et al. (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. Cell *116*, 499–509.

Di Cello, F., Cope, L., Li, H., Jeschke, J., Wang, W., Baylin, S.B., and Zahnow, C.A. (2013). Methylation of the claudin 1 promoter is associated with loss of expression in estrogen receptor positive breast cancer. PLoS One *8*, e68630.

Dang, C.V. (2012). MYC on the path to cancer. Cell *149*, 22–35.

DiNardo, D.N., Butcher, D.T., Robinson, D.P., Archer, T.K., and Rodenhiser, D.I. (2001). Functional analysis of CpG methylation in the BRCA1 promoter region. Oncogene *20*, 5331–5340.

Doose, G., Haake, A., Bernhart, S.H., Lopez, C., Duggimpudi, S., Wojciech, F., Bergmann, A.K., Borkhardt, A., Burkhardt, B., Claviez, A., et al. (2015). MINCR is a MYC-induced lncRNA able to modulate MYC's transcriptional network in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA *112*, E5261–E5270.

Du, Z., Fei, T., Verhaak, R.G., Su, Z., Zhang, Y., Brown, M., Chen, Y., and Liu, X.S. (2013). Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. Nat. Struct. Mol. Biol. *20*, 908–913.

Du, Z., Sun, T., Hacisuleyman, E., Fei, T., Wang, X., Brown, M., Rinn, J.L., Lee, M.G., Chen, Y., Kantoff, P.W., and Liu, X.S. (2016). Integrative analyses reveal a long noncoding RNA-mediated sponge regulatory network in prostate cancer. Nat. Commun. *7*, 10982.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Engel, J.R., Smallwood, A., Harper, A., Higgins, M.J., Oshimura, M., Reik, W., Schofield, P.N., and Maher, E.R. (2000). Epigenotype-phenotype correlations in Beckwith-Wiedemann syndrome. J. Med. Genet. *37*, 921–926.

Fernandez, P.C., Frank, S.R., Wang, L., Schroeder, M., Liu, S., Greene, J., Cocito, A., and Amati, B. (2003). Genomic targets of the human c-Myc protein. Genes Dev. *17*, 1115–1129.

Gartel, A.L., and Radhakrishnan, S.K. (2005). Lost in transcription: p21 repression, mechanisms, and consequences. Cancer Res. *65*, 3980–3985.

Gartel, A.L., Ye, X., Goufman, E., Shianov, P., Hay, N., Najmabadi, F., and Tyner, A.L. (2001). Myc represses the p21(WAF1/CIP1) promoter and interacts with Sp1/Sp3. Proc. Natl. Acad. Sci. USA *98*, 4510–4515.

Guo, H., Ahmed, M., Zhang, F., Yao, C.Q., Li, S., Liang, Y., Hua, J., Soares, F., Sun, Y., Langstein, J., et al. (2016). Modulation of long noncoding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. Nat. Genet. *48*, 1142–1150.

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., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071–1076.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223–227.

Hancock, A.L., Brown, K.W., Moorwood, K., Moon, H., Holmgren, C., Mardikar, S.H., Dallosso, A.R., Klenova, E., Loukinov, D., Ohlsson, R., et al. (2007). A CTCF-binding silencer regulates the imprinted genes AWT1 and WT1-AS and exhibits sequential epigenetic defects during Wilms' tumourigenesis. Hum. Mol. Genet. *16*, 343–354.

Hu, X., Feng, Y., Zhang, D., Zhao, S.D., Hu, Z., Greshock, J., Zhang, Y., Yang, L., Zhong, X., Wang, L.P., et al. (2014). A functional genomic approach identifies FAL1 as an oncogenic long noncoding RNA that associates with BMI1 and represses p21 expression in cancer. Cancer Cell *26*, 344–357.

Hung, C.L., Wang, L.Y., Yu, Y.L., Chen, H.W., Srivastava, S., Petrovics, G., and Kung, H.J. (2014). A long noncoding RNA connects c-Myc to tumor metabolism. Proc. Natl. Acad. Sci. USA *111*, 18697–18702.

Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Goncalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). A landscape of pharmacogenomic interactions in cancer. Cell *166*, 740–754.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat. Genet. *41*, 178–186.

Iyer, M.K., Niknafs, Y.S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T.R., Prensner, J.R., Evans, J.R., Zhao, S., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. Nat. Genet. *47*, 199–208.

Izumi, H., Molander, C., Penn, L.Z., Ishisaki, A., Kohno, K., and Funa, K. (2001). Mechanism for the transcriptional repression by c-Myc on PDGF beta-receptor. J. Cell Sci. *114*, 1533–1544.

Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. Nat. Rev. Genet. *3*, 415–428.

Kress, T.R., Sabo, A., and Amati, B. (2015). MYC: connecting selective transcriptional control to global RNA production. Nat. Rev. Cancer *15*, 593–607.

Lee, B.K., Bhinge, A.A., Battenhouse, A., McDaniell, R.M., Liu, Z., Song, L., Ni, Y., Birney, E., Lieb, J.D., Furey, T.S., et al. (2012). Cell-type specific and combinatorial usage of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. Genome Res. *22*, 9–24.

Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y., and Pietenpol, J.A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J. Clin. Invest. *121*, 2750–2767.

von der Lehr, N., Johansson, S., Wu, S., Bahram, F., Castell, A., Cetinkaya, C., Hydbring, P., Weidung, I., Nakayama, K., Nakayama, K.I., et al. (2003). The F-box protein Skp2 participates in c-Myc proteosomal degradation and acts as a cofactor for c-Myc-regulated transcription. Mol. Cell *11*, 1189–1200.

Leucci, E., Vendramin, R., Spinazzi, M., Laurette, P., Fiers, M., Wouters, J., Radaelli, E., Eyckerman, S., Leonelli, C., Vanderheyden, K., et al. (2016). Melanoma addiction to the long non-coding RNA SAMMSON. Nature *531*, 518–522.

Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A., and Dewey, C.N. (2010a). RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics *26*, 493–500.

Li, Y., Zou, L., Li, Q., Haibe-Kains, B., Tian, R., Li, Y., Desmedt, C., Sotiriou, C., Szallasi, Z., Iglehart, J.D., et al. (2010b). Amplification of LAPTM4B and YWHAZ contributes to chemotherapy resistance and recurrence of breast cancer. Nat Med *16*, 214–218.

Li, J.P., Liu, L.H., Li, J., Chen, Y., Jiang, X.W., Ouyang, Y.R., Liu, Y.Q., Zhong, H., Li, H., and Xiao, T. (2013). Microarray expression profile of long noncoding RNAs in human osteosarcoma. Biochem. Biophys. Res. Commun. *433*, 200–206.

Li, H., Chiappinelli, K.B., Guzzetta, A.A., Easwaran, H., Yen, R.W., Vatapalli, R., Topper, M.J., Luo, J., Connolly, R.M., Azad, N.S., et al. (2014). Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. Oncotarget *5*, 587–598.

Li, Z., Van Calcar, S., Qu, C., Cavenee, W.K., Zhang, M.Q., and Ren, B. (2003). A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. Proc. Natl. Acad. Sci. USA *100*, 8164–8169.

Luscher, B. (2001). Function and regulation of the transcription factors of the Myc/Max/Mad network. Gene *277*, 1–14.

Mei, S., Qin, Q., Wu, Q., Sun, H., Zheng, R., Zang, C., Zhu, M., Wu, J., Shi, X., Taing, L., et al. (2017). Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. Nucleic Acids Res. *45*, D658–D662.

Nelson, J.D., Denisenko, O., and Bomsztyk, K. (2006). Protocol for the fast chromatin immunoprecipitation (ChIP) method. Nat. Protoc. *1*, 179–185.

Niu, Y., Xu, M., Slagle, B.L., Huang, H., Li, S., Guo, G.L., Shi, G., Qin, W., and Xie, W. (2017). Farnesoid X receptor ablation sensitizes mice to hepatitis b virus X protein-induced hepatocarcinogenesis. Hepatology *65*, 893–906.

Noushmehr, H., Weisenberger, D.J., Diefes, K., Phillips, H.S., Pujara, K., Berman, B.P., Pan, F., Pelloski, C.E., Sulman, E.P., Bhat, K.P., et al. (2010). Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell *17*, 510–522.

Peukert, K., Staller, P., Schneider, A., Carmichael, G., Hanel, F., and Eilers, M. (1997). An alternative pathway for gene regulation by Myc. EMBO J. *16*, 5672–5686.

Prensner, J.R., and Chinnaiyan, A.M. (2011). The emergence of lncRNAs in cancer biology. Cancer Discov. *1*, 391–407.

Roy, A.L., Carruthers, C., Gutjahr, T., and Roeder, R.G. (1993). Direct role for Myc in transcription initiation mediated by interactions with TFII-I. Nature *365*, 359–361.

Ruan, W., Wang, P., Feng, S., Xue, Y., and Li, Y. (2016). Long non-coding RNA small nucleolar RNA host gene 12 (SNHG12) promotes cell proliferation and migration by upregulating angiomotin gene expression in human osteosarcoma cells. Tumour Biol. *37*, 4065–4073.

Sanchez, Y., Segura, V., Marin-Bejar, O., Athie, A., Marchese, F.P., Gonzalez, J., Bujanda, L., Guo, S., Matheu, A., and Huarte, M. (2014). Genome-wide analysis of the human p53 transcriptional network unveils a lncRNA tumour suppressor signature. Nat. Commun. *5*, 5812.

Schmitt, A.M., and Chang, H.Y. (2016). Long noncoding RNAs in cancer pathways. Cancer Cell *29*, 452–463.

Schmitt, A.M., Garcia, J.T., Hung, T., Flynn, R.A., Shen, Y., Qu, K., Payumo, A.Y., Peres-da-Silva, A., Broz, D.K., Baum, R., et al. (2016). An inducible long noncoding RNA amplifies DNA damage signaling. Nat. Genet. *48*, 1370–1376.

Scott, R.H., Douglas, J., Baskcomb, L., Huxter, N., Barker, K., Hanks, S., Craft, A., Gerrard, M., Kohler, J.A., Levitt, G.A., et al. (2008). Constitutional 11p15 abnormalities, including heritable imprinting center mutations, cause nonsyndromic Wilms tumor. Nat. Genet. *40*, 1329–1334.

Shen, H., and Laird, P.W. (2013). Interplay between the cancer genome and epigenome. Cell *153*, 38–55.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. *15*, 272–286.

Shrivastava, A., Saleque, S., Kalpana, G.V., Artandi, S., Goff, S.P., and Calame, K. (1993). Inhibition of transcriptional regulator Yin-Yang-1 by association with c-Myc. Science *262*, 1889–1892.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689–693.

Tseng, Y.Y., Moriarity, B.S., Gong, W., Akiyama, R., Tiwari, A., Kawakami, H., Ronning, P., Reuland, B., Guenther, K., Beadnell, T.C., et al. (2014). PVT1 dependence in cancer with MYC copy-number increase. Nature *512*, 82–86.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. Science *339*, 1546–1558.

Wagner, G.P., Kin, K., and Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. *131*, 281–285.

Wijnen, M., Alders, M., Zwaan, C.M., Wagner, A., and van den Heuvel-Eibrink, M.M. (2012). KCNQ1OT1 hypomethylation: a novel disguised genetic predisposition in sporadic pediatric adrenocortical tumors? Pediatr. Blood Cancer *59*, 565–566.

Wu, S.C., Kallin, E.M., and Zhang, Y. (2010). Role of H3K27 methylation in the regulation of lncRNA expression. Cell Res. *20*, 1109–1116.

Xiang, J.F., Yin, Q.F., Chen, T., Zhang, Y., Zhang, X.O., Wu, Z., Zhang, S., Wang, H.B., Ge, J., Lu, X., et al. (2014). Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. Cell Res. *24*, 513–531.

Yan, X., Hu, Z., Feng, Y., Hu, X., Yuan, J., Zhao, S.D., Zhang, Y., Yang, L., Shan, W., He, Q., et al. (2015). Comprehensive genomic characterization of long non-coding RNAs across human cancers. Cancer Cell *28*, 529–540.

Yang, F., Xue, X., Zheng, L., Bi, J., Zhou, Y., Zhi, K., Gu, Y., and Fang, G. (2014). Long non-coding RNA GHET1 promotes gastric carcinoma cell proliferation by increasing c-Myc mRNA stability. FEBS J. *281*, 802–813.

Zeller, K.I., Zhao, X., Lee, C.W., Chiu, K.P., Yao, F., Yustein, J.T., Ooi, H.S., Orlov, Y.L., Shahab, A., Yong, H.C., et al. (2006). Global mapping of c-Myc binding sites and target gene networks in human B cells. Proc. Natl. Acad. Sci. USA *103*, 17834–17839.

Zhou, Y., Zhang, X., and Klibanski, A. (2012). MEG3 noncoding RNA: a tumor suppressor. J. Mol. Endocrinol. *48*, R45–R53.

Zhu, S., Li, W., Liu, J., Chen, C.H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J., et al. (2016). Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. Nat. Biotechnol. *34*, 1279–1286.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
SNRP70	Abcam	Cat# Ab83306; RRID: AB_10673827
MAX	Novus	Cat# NBP1-49963; RRID: AB_10012153
GAPDH	Santa Cruz	Cat# sc-25778; RRID: AB_10167668
Cyclin A2	Santa Cruz	Cat# sc-596; RRID: AB_631330
MYC	Santa Cruz	Cat# sc-789; RRID: AB_631274
MAX	Santa Cruz	Cat# sc-764; RRID: AB_631276
MYC	Cell signaling	Cat# 13987; RRID: AB_2631168
MYC	Cell signaling	Cat# 9402; RRID: AB_10693752
p21	Cell signaling	Cat# 2947; RRID: AB_823586
CDC20	Cell signaling	Cat# 14866; RRID: AB_2715567
CDC45	Cell signaling	Cat# 11881; RRID: AB_2715569
FLAG	Cell signaling	Cat# 14793; RRID: AB_2572291
Normal Rabbit IgG	Cell signaling	Cat# 2729; RRID: AB_2617119
β-actin	Sigma-Aldrich	Cat# A5441; RRID: AB_476744
Anti-FLAG M2 affinity gel	Sigma-Aldrich	Cat# A2220; RRID: AB_10063035
Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	ThermoFisher	Cat# 31430; RRID: AB_228307
Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	ThermoFisher	Cat# 31460; RRID: AB_228341
Rabbit TrueBlot: Anti-Rabbit IgG HRP	Rockland	Cat# 18-8816-33; RRID: AB_2610848
Mouse TrueBlot ULTRA: Anti-Mouse IgG HRP	Rockland	Cat# 18-8817-33; RRID: AB_2610851
Chemicals, Peptides, and Recombinant Proteins		
Lipofectamine 2000 Reagent	ThermoFisher	Cat# 11668019
Lipofectamine RNAiMAX Reagent	ThermoFisher	Cat# 13778150
Lipofectamine 3000 Reagent	ThermoFisher	Cat# L3000015
T7 RNA Polymerase	Roche	Cat# 10881775001
PfuUltra II Fusion HotStart DNA Polymerase	Agilent Technologies	Cat# 600674
RNase Cocktail Enzyme Mix	ThermoFisher	Cat# AM2286
Proteinase K	NEB	Cat# P8107S
RQ1 DNase	Promega	Cat# M198A
BD Pharmingen Stain Buffer	BD Biosciences	Cat# 554656
BD Pharmingen PI/RNase Staining Buffer	BD Biosciences	Cat# 550825
Pierce Protein A/G Agarose	ThermoFisher	Cat# 20421
Dynabeads MyOne Streptavidin C1	ThermoFisher	Cat# 65001
Protease Inhibitor Cocktail	Sigma-Aldrich	Cat# P8340
Halt Protease& Phosphatase Single-Use Inhibitor Cocktail	ThermoFisher	Cat# 78442
Puromycin	Fisher BioReagents	Cat# BP2956-100
Hexadimethrine bromide	Sigma-Aldrich	Cat# H9268
Decitabine	LC Laboratories	Cat# D-3899
30% Acrylamide/Bis Solution 29:1	Bio-Rad	Cat# 161-0156
10x Tris/Glycine/SDS Buffer	Bio-Rad	Cat# 161-0772
Immun-Blot PVDF Membrane	Bio-Rad	Cat# 162-0177
Blue X-Ray Film	Phenix	Cat# F-BX810
Tween 20	Fisher BioReagents	Cat# BP337-100
Triton X-100	Bio-Rad	Cat# 161-0407
Agarose	Bio-Rad	Cat# 161-3102

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NonFat Dry Milk	Lab Scientific	Cat# M0841
Nonidet-P40	US Biological	Cat# N3500
Bovine Serum Albumin	Santa Cruz	Cat# sc-2323
Bovine Serum Albumin	Sigma-Aldrich	Cat# A2153
NuSieve GTG Agarose	Lonza	Cat# 50081
Recombinant Human MYC protein	Abcam	Cat# Ab84132
Critical Commercial Assays		
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems	Cat# 4368813
Biotin RNA Labeling Mix	Roche	Cat# 11685597910
Power SYBR Green PCR Master Mix	Applied Biosystems	Cat# 4367660
ECL Western Blotting Substrate	ThermoFisher	Cat# 32106
PARIS Kit	ThermoFisher	Cat# AM1921
BCA Protein Assay Kit	ThermoFisher	Cat# 23225
FirstChoice RLM-RACE Kit	ThermoFisher	Cat# AM1700
QuickChange II XL Site-Direct Mutagenesis Kit	Agilent Technologies	Cat# 200522
CellTiter 96 Non-Radioactive Cell Proliferation Assay	Promega	Cat# G4100
TOPO TA Cloning Kit	ThermoFisher	Cat# 1715582
AmpliTaq Gold 360 Master Mix	Applied Biosystems	Cat# 4398881
Deposited Data		
RNA-seq based gene expression in MCF-7 cells after siRNA-mediated knockdown of *EPIC1*	This paper	GSE98538
DNA methylation data of breast cancer cells	(Li et al., 2014)	GSE57342
DNA methylation data of breast cancer cells	(Di Cello et al., 2013)	GSE44837
Gene expression profile of breast cancer patients	https://www.ncbi.nlm.nih.gov/geo	GSE20711
Gene expression profile of breast cancer patients	https://www.ncbi.nlm.nih.gov/geo	GSE21653
Gene expression profile of breast cancer patients	https://www.ncbi.nlm.nih.gov/geo	GSE17907
Gene expression profile of breast cancer patients	https://www.ncbi.nlm.nih.gov/geo	GSE20685
Gene expression profile of breast cancer patients	https://www.ncbi.nlm.nih.gov/geo	GSE16446
Gene expression profile of breast cancer patients	(Li et al., 2010b)	GSE19615
Experimental Models: Cell Lines		
MCF10A	ATCC	Cat# CRL-10317; RRID: CVCL_10317
BT-20	ATCC	Cat# HTB-19; RRID: CVCL_0178
BT-474	ATCC	Cat# HTB-20; RRID: CVCL_0179
HCC1937	ATCC	Cat# CRL-2336; RRID: CVCL_0290
Hs578T	ATCC	Cat# HTB-126; RRID: CVCL_0332
MDA-MB-231 (MB231)	ATCC	Cat# HTB-26; RRID: CVCL_0062
MDA-MB-361 (MB361)	ATCC	Cat# HTB-27; RRID: CVCL_0620
MDA-MB-468 (MB468)	ATCC	Cat# HTB-132; RRID: CVCL_0419
MCF-7	ATCC	Cat# HTB-22; RRID: CVCL_0031
T-47D	ATCC	Cat# HTB-133; RRID: CVCL_0553
ZR-75-1	ATCC	Cat# CRL-1500; RRID: CVCL_0588
A2780	ECACC	Cat# 93112519; RRID: CVCL_0134
A2780cis	ECACC	Cat# 93112517; RRID: CVCL_1942
IGR-OV-1	NIH/NCI	Cat# IGR-OV1; RRID: CVCL_1304
NIH:OVCAR-3	ATCC	Cat# HTB-161; RRID: CVCL_0465
OVCAR-4	NIH/NCI	Cat# OVCAR-4; RRID: CVCL_1627
OVCAR-8	NIH/NCI	Cat# OVCAR-8; RRID: CVCL_1629
SK-OV-3	ATCC	Cat# HTB-77; RRID: CVCL_0532

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
293T	ATCC	Cat# CRL-3216; RRID: CVCL_0063
E. coli DH5α Competent Cells	ThermoFisher	Cat# 18265017
E. coli Stbl3 Competent Cells	ThermoFisher	Cat# C737303
Experimental Models: Organisms/Strains		
Athymic nude Mice	Charles River	N/A
Oligonucleotides		
siRNA, LNA, and shRNA target sequence, see Table S5	This paper	N/A
Primers for qRT-PCR and ChIP-PCR, see Table S5	This paper	N/A
Primers for plasmid constructs, see Table S5	This paper	N/A
Primers for *in vitro* transcribed RNA, see Table S5	This paper	N/A
Recombinant DNA		
pBABE Puro	Addgene	Cat# 1764
pLnxEXP	Addgene	Cat# 64865
pBABE-lnc	This paper	N/A
pBABE-lnc EPIC1 v1	This paper	N/A
pBABE-lnc EPIC1 v2	This paper	N/A
pBABE-lnc EPIC1 v3	This paper	N/A
psPAX2	Addgene	Cat# 12260
pVSV-G	Addgene	Cat# 8454
pLKO.1 TRC Cloning Vector	Addgene	Cat# 10878
pLKO.1 shCtrl	This paper	N/A
pLKO.1 shEPIC1 1#	This paper	N/A
pLKO.1 shEPIC1 2#	This paper	N/A
pCDH-CMV-MCS-EF1-Puro	System Biosciences	Cat# CD510B-1
pCDH EPIC1 v1	This paper	N/A
pCDH EPIC1 v1 Δ1-60nt	This paper	N/A
pCDH EPIC1 v1 Δ61-120nt	This paper	N/A
pCDH EPIC1 v1 Δ121-180nt	This paper	N/A
pCDH EPIC1 v1 Δ181-240nt	This paper	N/A
pCDH EPIC1 v1 Δ241-300nt	This paper	N/A
pCDH EPIC1 v1 Δ301-358nt	This paper	N/A
pCDH EPIC1 v1 Δ129-283nt	This paper	N/A
pCDH EPIC1 v1 LNA_R	This paper	N/A
pCDH EPIC1 v1 Δ129-283nt LNA_R	This paper	N/A
WWP-Luc (p21/WAF1 promoter)	Addgene	Cat# 16451
pGL3 Basic	Promega	Cat# E1751
EPIC1-Luc	This paper	N/A
CCNA2-Luc	This paper	N/A
β-Gal	(Niu et al., 2017)	N/A
pCMV6-XL5-MYC	OriGene	Cat# SC112715
pCMV-Tag2B (Flag)	Agilent Technologies	Cat# 211172
pCMV-Flag-MYC	This paper	N/A
pCMV-Flag-MYC N220	This paper	N/A
pCMV-Flag-MYC N366	This paper	N/A
pCMV-Flag-MYC ΔN147	This paper	N/A
pCMV-Flag-MYC Δ148-220 (ΔEPIC1)	This paper	N/A
pCMV-Flag-MYC ΔN220	This paper	N/A
pLVX-IRES-Hygro	Clontech	Cat# 632185
pLVX-HA-MAX Hygro	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pLVX-HA-MYC Hygro	This paper	N/A
pLVX-HA-MYC Δ148-220 (ΔEPIC1) Hygro	This paper	N/A
Software and Algorithms		
GraphPad Prism	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
FlowJo	FLOWJO, LLC	https://www.flowjo.com/solutions/flowjo
ImageJ	ImageJ	https://imagej.nih.gov/ij/download.html
R	R	https://www.r-project.org/about.html
STAR	STAR	https://github.com/alexdobin/STAR
RSEM	RSEM	https://deweylab.github.io/RSEM/
GSEA	GSEA	http://software.broadinstitute.org/gsea/index.jsp
ggtern	R package	https://cran.r-project.org/web/packages/ggtern/
survival	R package	https://cran.r-project.org/web/packages/survival/index.html
ggplot2	R package	http://ggplot2.org/
CISTROM	Cistrome Project	http://www.cistrome.org/Cistrome/Cistrome_Project.html

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Da Yang (dyang@pitt.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Culture, RNA Interference, LNA Transfection, and Plasmid Transfection

Human breast epithelial cell line, MCF10A, and human breast cancer cell lines, BT-20, BT-474, HCC1937, Hs578T, MCF-7, MDA-MB-231 (MB231), MDA-MB-361 (MB361), MDA-MB-468 (MB468), T-47D, and ZR-75-1, and human ovarian cancer cell lines, SK-OV-3, and NIH: OVCAR-3, and human pancreatic cancer cell lines, AsPC-1, BxPC-3, and PANC-1, and human prostate cancer cell lines, DU 145, and PC-3, and human leukemia cell line K562, and human lung cancer cell line A549, and human cervical cancer cell line HeLa, and human liver cancer cell line Hep G2, and human embryonic kidney (HEK) 293T cells were purchased from American Type Culture Collection (ATCC) and cultured as suggested by ATCC's guidelines. Human ovarian cancer cell lines, IGR-OV-1, OVCAR-4, and OVCAR-8 were purchased from NIH/NCI and kept in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin, and 1% streptomycin. The A2780 human ovarian cancer cell line and the cisplatin resistant version of the cell line, A2780cis, were obtained from the European Collection of Cell Cultures (ECACC), supplied by Sigma-Aldrich, and cultured in RPMI 1640 medium supplemented with 2 mM glutamine, 10% FBS, 1% penicillin, and 1% streptomycin; A2780cis cells were also supplemented with 1 μM cisplatin. Human pancreatic duct epithelial cell line (HPDE), and phoenix cells were kindly provided by Dr. Wen Xie (Department of Pharmaceutical Sciences, University of Pittsburgh), and HPDE cells were maintained in Keratinocyte-SFM medium supplemented with human recombinant epidermal growth factor and bovine pituitary extract (ThermoFisher, #17005042) and phoenix cells were maintained in DMEM supplemented with 10% FBS, 1% penicillin, and 1% streptomycin.

For RNA interference, cells were transfected with 40 nM siRNA targeting *EPIC1*, *MYC*, or a control siRNA using Lipofectamine RNAiMAX (ThermoFisher, #13778150) per the manufacturer's instructions. Total RNA was isolated 72 hr later for real-time PCR analysis. The siRNA sequences are listed in Table S5.

For LNA transfection, cells were transfected with 40 nM LNA oligos targeting *EPIC1*, and a scramble control using Lipofectamine™ RNAiMAX per the guidelines. The LNA oligos were designed and synthesized from Exiqon, and detailed sequences are listed in Table S5.

For plasmid transfection, cells were transfected with plasmid using Lipofectamine™ 2000 (ThermoFisher, #11668019) or Lipofectamine™ 3000 (ThermoFisher, #L3000015) as suggested approaches.

In Vivo Xenograft Model

Briefly, 5- to 6-week-old female athymic nude mice (Charles River) were used for the xenograft model. MCF-7 cells stably expressing shCtrl and sh*EPIC1* were trypsinized and washed twice with sterilized PBS, and then, 0.2 ml of PBS containing 5 x 10^6 cells was subcutaneously inoculated into the flanks of the mice. Mice were monitored twice every week for tumor growth, and tumor size was measured using a caliper. Tumor volume in mm^3 was calculated using the formula: Tumor volume = 0.5 x $(width)^2$ × length. Eight weeks after inoculation, mice were sacrificed in keeping with the policy for the humane treatment of tumor-bearing animals. All animal studies were performed in accordance with the institutional guidelines, and the experiments followed the protocols approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Pittsburgh.

METHODS DETAILS

Data Collection

DNA methylation, PCG expression, whole-exome mutation and GISTIC copy number alteration data were downloaded from TCGA Pan-Cancer project (Data Freeze 1.3). The lncRNA annotation was downloaded from GENCODE (V22, GRCh38). There were 7,656 intergenic, 5,565 antisense, and 920 sense intronic lncRNAs. H3K4me3 and H3K27ac ChIP-seq data for seven cell lines were downloaded from the UCSC genome browser: Integrated Regulation from ENCODE Tracks. DNA methylation data for breast cancer cell lines were downloaded from GSE57342 (Li et al., 2014) and GSE44837 (Di Cello et al., 2013).

RNA-seq data from 781 cancer cell lines in the CCLE database were downloaded from Expression Atlas (E-MTAB-2770). HM450 DNA methylation profile of 1,028 cancer cells lines form COSMIC database (Iorio et al., 2016). There are 455 cells which have both HM450 DNA methylation and RNA-seq data. The BAM files of RNA-seq of 939 breast cancer tumors were downloaded from Cancer Genomics Hub.

Mapping the Probes to GENCODE Genes

The genomic coordinates of HM450 probes based on GRCh37 were first transferred to genomic coordinates in GRCh38 using LiftOver (UCSC genome browser). We then searched the nearest TSS of PCG and lncRNA for each probe based on GENCODE V22 annotation. In this way, we defined: (1) the PCG probes, located in the PCG promoter region (+/- 3 kb from the TSS); (2) the lncRNA probes, located in the lncRNA promoter region; (3) the shared probes, located in both the PCG and lncRNA promoter regions; and (4) the non-probes, which are not located in any promoter regions (Figure S1B).

DNA Methylation Dysregulation Pattern Analysis in Cancers

DNA methylation dysregulation in cancers showed a different beta value pattern in lncRNA promoter and protein-coding promoter regions. To evaluate the statistical significance of the difference between methylation in lncRNA and PCG promoter regions, we permuted the annotation for each probe 10,000 times to generate an experimental distribution of DNA methylation change. Through comparison with the experimental distribution, an empirical p value could be calculated. Finally, the weighted two-dimensional kernel density estimation R function *kde2d.weighted* (package: ggtern) was used to measure the distribution of hypomethylation or hypermethylation according to the distance to promoters of lncRNA and PCGs.

MiTranscriptome Data Renormalization

Recent reports have revealed that highly expressed genes affect the normalization scale much more and cause a bias against low-expression genes such as lncRNAs (Li et al., 2010a; Wagner et al., 2012). To precisely evaluate the alteration of lncRNA expression in tumors, we renormalized the MiTranscriptome profile using a method similar to that described in S. Anders et al. (Anders and Huber, 2010). Specifically, a scaling factor for each sample was calculated as the median of the expression ratio to a pseudo-reference sample for each gene. The pseudo-reference sample was computed as the median expression level across all samples for that gene.

The formula to calculate the i-th sample's scaling factor:

$$scale_i = median_{j=1\ldots n}\left(\frac{E_{ij}}{median_{i=1\ldots m}(E_{ij})}\right)$$

where E indicates the expression profile, which has m samples and n genes. The denominator of the formula can be interpreted as j-th gene expression level of the pseudo-reference sample.

Characterization of the lncRNA Landscape

We used a strategy similar as described in TCGA Glioblastoma project (Brennan et al., 2013) to characterize the epigenetic lncRNA landscape in each cancer type, which has successfully generated a patient-centric matrix for PCGs in glioblastoma using an Infinium HumanMethylation27 microarray. We adapted the strategy to accommodate for the lncRNA genes and the HumanMethylation450 microarray. Specifically, we first identified lncRNA and HM450 probe pairs in which the probe located at the lncRNA's promoter region as described previously. Then Spearman correlation coefficients (Rho) between the methylation alteration and gene expression for each lncRNA and probe pair were calculated for each cancer type. The probe with highest coefficient was selected for the lncRNA if multiple probes annotated to same gene promoter to capture the most variable and correlated probe for each gene. This

procedure reduced the number of CpG probes from N:1 to 1:1. Next, we assigned discrete categories based on the Spearman correlation coefficient according to the following criteria:

1. Strongly negatively correlated (SNC) when the rho value was less than -0.5;
2. Weakly negatively correlated (WNC) when the rho value was between -0.5 and -0.25;
3. No negative correlation (NNC) when the rho value was greater than -0.25.

Next, we assigned samples to either the 30th (T30 or N30) or 70th (T70 or N70) percentile based on the observed beta value across tumor (T) and normal (N) samples. For a cancer type with less than 30 normal samples, we randomly selected 24 normal samples from each of the three different normal tissues (72 samples in total). The three different normal tissues selected for this analysis were generated by TCGA for breast (BRCA), kidney (KIRC) and lung (LUSC) tumor studies. We finally scored each lncRNA gene per cancer type per tissue type (tumor and normal) according to the following rules:

1.If percentile 70 < 0.25, we score it as CUN or CUT (constitutively unmethylated in normal or tumor tissue);
2.If percentile 20 > 0.75, we score it as CMN or CMT (constitutively methylated in normal or tumor tissue);
3.If percentile 20 > 0.25 and percentile 70 < 0.75, we score it as IMN or IMT (intermediately methylated in normal or tumor tissue);
4.If it did not fall into any of the above categories, it was scored VMN or VMT (variably methylated in normal or tumor tissue).

Next, we assigned a 'call' and a confidence 'score' for each of the possible combinations (48) [3 (SNC, WNC, NNC) x 4 (CUN, CMN, VMN, IMN) x 4 (CUT, CMT, VMT, IMT)] per platform, as shown in Table S5. The methylation calls are as follows:

EA: Epigenetically activated
ES: Epigenetically silenced
UC: No Change

Methylation class confidence scores varied from EAH (epigenetic activation with high confidence), EAL (epigenetic activation with low confidence), NC (no change), ESL (epigenetic silencing with low confidence) and ESH (epigenetic silencing with high confidence) here. In this way, we generated a Methylation Patient-Centric Table of DNA methylation calls for each sample per lncRNA in 20 cancer types, and calculated the percentage of four types' methylation status for each lncRNA in each cancer type.

The lncRNAs were ranked by summarized weighted alteration percentages among all the cancer types. Specifically, we give the EAH percentage with weight 2, EAL percentage with weight 1, UC percentage with weight 0, ESL percentage with weight -1, and ESH percentage with weight -2. The summarized weighted percentages of each lncRNA was used as a rank score. Generally, lncRNAs with consistent EA status in multiple cancer types would get a higher score, and the lncRNAs with consistent ES status in multiple cancer types would show a lower score.

Gene Set Enrichment Analysis (GSEA)

To interpret the function of regulated genes after *EPIC1* siRNA treatment, GSEA (version 2.2.0) (Subramanian et al., 2005) was performed using the 50 cancer hallmark gene sets and a gene log2-fold change. To identify the pathways that are correlated with *EPIC1* expression in tumor samples, we performed a similar GSEA for each cancer type in TCGA dataset. In this analysis, GSEA was performed on the ranked PCG list based on the Spearmen's correlation coefficient with *EPIC1* expression.

RNA-seq Data Analysis

We developed a STAR-RSEM pipeline, which was revised from the ENCODE RNA-seq analysis pipeline. We used this pipeline to profile TCGA breast cancer and CCLE breast cancer cell line RNA-seq data, and the RNA-seq data of MCF-7 cells after *EPIC1* knockdown. To transfer the bam file to fastq, we used Picard-tools SamToFastq module. FastQC was used to check the sequencing quality. The RNA-seq data can be downloaded from GEO (GSE98538).

Association Analysis between lncRNA Epigenetic Landscape and Protein-Coding Gene Alteration

Somatic mutations and copy number alterations in 32 cancer types were obtained from TCGA Pan-Cancer project (http://cancergenome.nih.gov/tcga/). The somatic mutations were identified via the MC3 algorithm. The copy number alterations were called using the GISTIC algorithm. An alteration profile of 32 cancer types was constructed. The columns of the alteration profile represent the samples, and the rows represent the tumor genes. If a gene was detected with alterations (non-synonymous somatic mutation or SCNA) in a sample, we set the profile to 1. Otherwise, the profile was set to 0.

For each PCG-lncRNA pair (denoted as G_i and L_i), we calculated the probability $P_{(Gi,\ Li)}$ of observing at least the number of samples that simultaneously contain alterations in both G_1 and L_i at random according to Equation 1:

$$P_{(G_1,L_i)} = 1 - \sum_{k=0}^{a-1} \frac{\binom{a+b}{k}\binom{c+d}{a+c-k}}{\binom{n}{a+c}} \qquad \text{(Equation 1)}$$

where n is the total number of samples, a is the number of samples with alterations in both genes, b is the number of samples with alterations only in G_1, c is the number of samples with alterations only in L_i, and d is the number of samples without alterations in either

gene. The "hypergeometric test" p value was subjected to a Benjamini and Hochberg correction for multiple tests, and gene pairs with a FDR less than 0.05 were included in the following analysis.

Statistical and Clustering Analysis

Student's *t*-test, analysis of variance, chi-square, Wilcoxon rank-sum test, Fisher's exact test, Kaplan-Meier estimate, and Mantel-Cox survival analyses were performed using R 2.10.0. Significance was defined as $p < 0.05$. Benjamini-Hochberg multiple testing correction (Benjamini and Hochberg, 1995) was used to estimate the FDR when multiple testing correction was applied.

Integrating ChIP-Seq and RNA-Seq Data to Identify and Validate *EPIC1*-MYC Axis Target Gene

The genome-wide MYC protein binding sites were identified by applying Cistrome algorithm (Mei et al., 2017) on two biological replicates of MYC ChIP-seq assays of MCF-7 cells (Lee et al., 2012). We identified MYC targets that regulated by *EPIC1* based on two criteria: (1) at least one MYC binding peak falls within the TSS-proximal region (from 3 kb upstream to 500 bp downstream) of the gene; and (2) the gene is differentially expressed between the si*EPIC1* and control MCF-7 cells. The top targets of *EPIC1*-MYC axis were selected based on their significance of MYC binding signal, differential expression after *EPIC1* knockdown, and their roles in cell proliferation/cycle. For each target, primers were designed to target the MYC binding region, and detailed primer sequences are listed in Table S5. ChIP-qPCR was further performed to demonstrate whether *EPIC1* knockdown decreases the recruitment of MYC to its target promoter sites.

Antibodies

The following antibodies were used for immunoblotting: rabbit anti-SNRP70 (Abcam, #ab83306), rabbit anti-GAPDH (Santa Cruz, #sc-25778), rabbit anti-MYC (Cell Signaling, #13987), rabbit anti-p21 (Cell Signaling, #2947), rabbit anti-CDC20 (Cell Signaling, #14866), rabbit anti-FLAG (Cell Signaling, #14793), rabbit anti-CDC45 (Cell Signaling, #11881), rabbit anti-MAX (Novus, #NBP1-49963), mouse anti-Cyclin A2 (Santa Cruz, #sc-596), and mouse anti-β-actin (Sigma, #A5441). The following antibodies were used for co-immunoprecipitation (Co-IP), RNA immunoprecipitation (RIP) and chromatin immunoprecipitation (ChIP) analysis: rabbit anti-MYC (Santa Cruz, #sc-789), rabbit anti-MAX (Santa Cruz, #sc-764), rabbit anti-MYC (Cell Signaling, #9402), and normal rabbit IgG (Cell Signaling, #2729) as a negative control, and anti-FLAG M2 affinity gel (Sigma, #A2220).

Cell Fractionation, Cytoplasmic/Nuclear RNA Isolation

MCF-7, Hs578T, and T-47D cells were subjected to cytoplasmic and nuclear fractionation using a PARIS^{TM} kit (ThermoFisher, #AM1921), and total RNA was isolated from each fraction following the recommended protocol.

RNA Isolation and Quantitative Real-Time PCR (qRT-PCR) Assays

Total RNA was isolated from cultured cells using an RNeasy Mini kit (Qiagen, #74104) according to the manufacturer's instructions. cDNAs were synthesized from 0.5 μg of total RNA using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, #4368813). Real-time PCR was performed with *Power* SYBR Green PCR Master Mix (Applied Biosystems, #4367659) on a QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems). Relative gene expression was determined by ΔΔCt normalized to *GAPDH*. The primers used are listed in Table S5.

EPIC1 RNA Copy Number Analysis

Total RNA was isolated from 1 x 10^6 cells using an RNeasy Mini kit. The full-length of *EPIC1* RNA was *in vitro* transcribed with Ribonucleotide solution set (NEB, #N0450) and T7 RNA polymerase (Roche, #10881775001) using the PCR products as a template, treated with RNase-free DNase I (Promega, #M198A), and then isolated with the RNeasy Mini kit. cDNA was synthesized using 1μg of the total RNA or full-length of *EPIC1* RNA. Serial ten-fold dilutions (10^2 to 10^9 molecules per μl) of cDNA from *in vitro*-transcribed *EPIC1* RNA were used as a reference molecule for the standard curve calculation. Real-time PCR was performed as above.

Cloning, shRNA Construction, and Lentiviral Transduction

Full-length of *EPIC1* was identified and amplified from total RNAs of MCF-7 / T-47D cells by 5'RACE and 3'-RACE using FirstChoice RLM-RACE Kit (ThermoFisher, #AM1700). To construct retroviral *EPIC1* expression plasmids, PCR products containing the CMV-zsGreen1 portion of pLncEXP (Addgene plasmid # 64865) were inserted into a pBABE puro vector (Addgene, #1764), and the resulting construct was named as pBABE-lnc. Then full-length and truncated mutants of *EPIC1* were cloned into pBABE-lnc with *Age*I and *Xho*I enzymes or cloned into pCDH-CMV-MCS-EF1-Puro (System Biosciences, #CD510B-1) with *Xba*I and *Eco*RI enzymes. Full-length of Flag-tagged or HA-tagged MYC expression vectors were generated using a human MYC cDNA Clone (OriGene, #SC112715) as a DNA template. Full-length of HA-tagged MAX expression vector was generated using cDNA from MCF-7 cells as a template. The truncated or deletion mutants and LNA-resistant *EPIC1* expression vectors were constructed by using QuickChange II XL Site-Direct Mutagenesis Kit (Agilent Technologies, #200522). All constructs were confirmed by DNA sequencing at Genomics Research Core, University of Pittsburgh.

To construct stable *EPIC1*-expressing cells, pBABE-lnc and lnc-*EPIC1* plasmids were transfected into Phoenix cells to produce retrovirus, and viruses were collected 48 hr post-transfection. MCF-7 cells were infected for 24 hr with the retroviruses and selected with puromycin to establish stable *EPIC1*-expressing cells. Detailed sequences of primers used for cloning are listed in Table S5.

EPIC1 knockdown constructs were cloned by inserting oligos into a pLKO.1 TRC cloning vector (Addgene, #10878). The oligo sequences are listed in Table S5. To produce lentiviral particles, HEK 293T cells were seeded into one 6-cm Petri dish in DMEM with 10% FBS without antibiotics and incubated overnight to reach approximately 80% confluence before transfection. Transfection was performed using Lipofectamine 2000 Transfection Reagent according to the recommended protocol. Then, 3 μg of pLKO.1 shControl (shCtrl) or pLKO.1 sh*EPIC1* plasmid, 2.25 μg of psPAX2 (Addgene, #12260), and 0.75 μg of pVSV-G (Addgene, #8454) were used for each 6-cm petri dish. After transfection for 6 hr, the medium was changed with fresh DMEM containing 10% FBS, and the cells were incubated for another 48 hr. Culture medium containing the lentiviral particles was collected and filtered through a 0.45 μm filter to remove any remaining cells and debris. Target cells were infected for 24 hr with lentiviral particles in the presence of 8 μg/ml polybrene and screened with puromycin to establish stable cells.

Promoter Cloning and Reporter Assay

Using genomic DNA from MCF-7 cells as DNA templates, the promoter region of *CCNA2* ranging from -443 bp to +334 bp was amplified by PCR and inserted to pGL3 Basic vector (Promega, #E1751) with *Nhe*I and *Hin*dIII enzymes, named as CCNA2-Luc, and the promoter region of *EPIC1* ranging from -133 bp to +587 bp were inserted to pGL3 Basic vector with *Hin*dIII enzymes, named as *EPIC1*-Luc. WWP-Luc (p21/WAF1 promoter) was a gift from Bert Vogelstein (Addgene plasmid #16451). For plasmid methylation followed by the previous report (DiNardo et al., 2001), briefly, 20 μg of *EPIC1*-Luc were methylated using Methyltransferase (M. SssI, NEB, #M0226S) at 37°C for 12 hr, followed by subsequent inactivation of enzyme at 60°C for 20 min. Mock-methylated mixtures were also performed in the absence of the methylase and S-adenosyl methionine. The methylated and mock-methylated mixtures were purified using

QIAprep Spin Miniprep Kit (Qiagen, #27106) and the methylation status of the constructs was determined by HpaII digestion and 2% agarose gel eletrophoresis.

Cells were transiently transfected with un-methylated or methylated *EPIC1*-Luc reporter or a combination of either *EPIC1* siRNA, *MYC* siRNA, or a negative control siRNA with CCNA2-Luc or WWP-Luc constructs using Lipofectamine™ 2000, and β-Gal was used as an internal control. After 48 hr, the luciferase and β-Gal activities were detected as described (Niu et al., 2017) in a Wallac 1420 Victor^2 Microplate Reader (Perkin Elmer). The luciferase activities were normalized to the β-Gal activities. Data were shown as fold change over the control group.

Cell Proliferation and Cell Cycle Assay

Cells were seeded at 2,000 cells per well in 96-well culture plates, and MTT assays were performed with a CellTiter 96 Non-Radioactive Cell Proliferation Assay Kit (Promega, #G4100) following the manufacturer's guidelines. The absorbance value was measured at 570 nm using an xMark Microplate Spectrophotometer (Bio-Rad) with a reference wavelength of 630 nm.

For the cell cycle assay, cells were collected, rinsed with PBS, and fixed for a minimum of 2 hr by adding 70% ice-cold ethanol at -20°C. Cells were then sequentially washed once in PBS and BD Pharmingen stain buffer (BD Biosciences, #554656). Cell pellets were resuspended in 0.5 ml of BD Pharmingen PI/RNase staining buffer (BD Biosciences, #550825) and incubated for 15 min at room temperature (RT), and cells were immediately analyzed using an LSRFORTESSA X-20 flow cytometer (BD Biosciences). The data were analyzed with FlowJo software.

Soft Agar Colony Formation Assay

For each well, 2 ml of 0.6% NuSieve GTG agarose (Lonza, #50081) in culture medium was plated into 6-well plates as the bottom layer, and the agarose was allowed to solidify at RT. Then, 1 ml of cell mixture containing 10^4 cells in culture medium and a final concentration of 0.35% agarose was carefully plated on top of the bottom layer. The plates were incubated at 37°C and 5% CO_2 until colonies were formed, and cells were fed with 0.5 ml of cell culture medium every other week. After 2-3 weeks, colonies were stained using 0.005% crystal violet in 4% paraformaldehyde solution and counted.

RNA Immunoprecipitation (RIP)

RIP was performed as previously described with minor modifications (Tsai et al., 2010). Briefly, cultured cells were collected by trypsinization, washed once with cold PBS, and then treated with 0.3% formaldehyde in PBS for 10 min at 37°C. Then, 1.25 M glycine dissolved in PBS was added to a final concentration of 0.125 M, and the mixture continued to incubate for 5 min at RT. The cells were subsequently washed twice with cold PBS, and the pellets were resuspended in RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate, 0.5 mM DTT, 1 mM PMSF, and 1 x protease inhibitor cocktail (Sigma, #P8340)) and incubated on ice for 30 min with shaking. The cleared lysates were incubated for 4 h at 4°C with the corresponding antibodies. Pellets were washed twice in RIPA buffer, four times in 1 M RIPA buffer (50 mM Tris-HCl, pH 7.4, 1 M NaCl, 1 mM EDTA, 0.1% SDS, 1% NP-40, and 0.5% sodium deoxycholate), and then twice in RIPA buffer. The pellets were resuspended and treated with RIPA buffer containing proteinase K at 45°C for 45 min. Finally, RNA was isolated with TRIzol reagent.

RNA Pull-Down Assay

Biotin-labeled full-length and truncated fragments of *EPIC1* RNA were transcribed *in vitro* with a Biotin RNA Labeling Mix Kit (Roche, #11685597910) and T7 RNA polymerase (Roche, #10881775001) using the PCR products as a template, treated with RNase-free DNase I (Promega, #M198A), and then isolated with an RNeasy Mini kit. Biotinylated RNA was folded in RNA structure buffer

(10 mM Tris-HCl pH 7.0, 0.1 M KCl, 10 mM $MgCl_2$) at 90°C for 2 min, immediately put on ice for another 2 min, and then transferred to RT for 20 min to allow proper RNA secondary structure formation.

Cells were collected by trypsinization and washed twice with sterilized PBS. Cell pellets were resuspended in 2 ml of pre-chilled PBS, 2 ml of nuclear isolation buffer (1.28 M sucrose, 40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, and 4% Triton X-100) and 6 ml of sterilized DEPC-treated water and incubated on ice for 20 min with frequent vortexing. Nuclei were pelleted by centrifugation at 2,500 *g* for 15 min, washed once with 1 ml of nuclear isolation buffer, resuspended in RIP buffer (150 mM KCl, 25 mM Tris-HCl pH 7.4, 0.5 mM DTT, 0.5% NP-40, 1 mM PMSF, 1 x Superase-in, and 1 x protease inhibitor cocktail), and sheared on ice using a Dounce homogenizer with 15 to 20 strokes. After 1 mg of the cleared lysate was mixed with folded RNA in RIP buffer and incubated for 1 hr at RT, 60 μl of Dynabeads MyOne Streptavidin C1 magnetic beads (ThermoFisher, #65001) was added to each reaction, and the mixture was incubated for another 1 hr at RT. Beads were washed five times and boiled in 1 x SDS loading buffer, and the retrieved protein was analyzed using western blotting.

The *in vitro* binding assay of biotin-labeled *EPIC1* RNA and MYC protein was performed as previously described (Tsai et al., 2010). Briefly, 0.1 μg of biotinylated RNA was incubated with different amounts of recombinant human MYC protein (Abcam, #ab84132) for 1 hr at RT in 200 μl of binding buffer (50 mM Tris-HCl pH 7.9, 10% glycerol, 100 mM KCl, 5 mM $MgCl_2$, 10 mM β-ME, 0.1% NP-40, 1 mM PMSF, 1 x Superase-in, and 1 x protease inhibitor cocktail). Then, 30 μl of washed streptavidin-conjugated magnetic beads were added to each reaction, and the mixtures were incubated at RT for 30 min. Beads were washed five times and boiled in 1 x SDS loading buffer, and the retrieved protein was analyzed using western blotting.

Chromatin Immunoprecipitation (ChIP)

The ChIP assay was performed as previously described (Nelson et al., 2006). Briefly, 1 x 10^7 cells were cross-linked with a final concentration of 1.42% formaldehyde in growth medium for 15 min at RT, and cross-linking was quenched by the addition of glycine to a final concentration of 125 mM and incubation for 5 min at RT. Cells were rinsed twice with cold PBS, harvested in IP buffer (50 mM pH 7.5 Tris-HCl, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, and 1% Triton X-100) supplemented with 1 mM PMSF and 1 x protease inhibitor cocktail and sonicated to shear the chromatin to yield DNA fragment sizes of 0.5 to 1 kb. Samples were cleared by centrifuging at 12,000 *g* for 10 min at 4°C and preincubated for 1 hr with 40 μl of protein A/G agarose beads. A portion of the precleared samples was used as input DNA. Then, approximately 2 μg of MYC antibody or rabbit normal immunoglobulin (IgG) was added to the remainder of the samples and incubated for 1 hr at 4°C, 40 μl of protein A/G agarose beads (ThermoFisher, #20421) were added, and the mixture was incubated for 4 hr at 4°C. Beads were washed six times with cold IP buffer, and DNA was isolated with 10% Chelex following the suggested protocol; the total input DNA was also isolated. Quantification was performed using real-time PCR with SYBR Green Master Mix. Control IgG and input DNA signal values were used to normalize the values from the MYC ChIP to target genes. The primers for target genes and the negative control are listed in Table S5.

Co-Immunoprecipitation (Co-IP), Protein Isolation and Western Blotting

Co-IP was performed as following, briefly, cells were collected and lysed in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDT, 1% Triton X-100, PMSF freshly added to a final concentration of 1mM, and 1x protease inhibitor cocktail). After quantification using a BCA protein assay kit (ThermoFisher, #23225), 1 mg of total protein were used for Co-IP and incubated for overnight with 2 μg of anti-MYC, anti-MAX antibodies, and normal rabbit IgG as a negative IP control, respectively. The mixtures were incubated for another 2-4 hr with protein A/G agarose beads, and then beads were washed at least 4 times, and treated and boiled for 10 min with 1x SDS sample buffer (Bio-Rad, #161-0737).

Cell lysates were also treated with equal volume of 2x SDS sample buffer and resolved by SDS-PAGE under denaturing conditions and transferred onto PVDF membranes (Bio-Rad, #162-0177). The membranes were blocked with 5% non-fat milk (LabScientific, #M0841) in 1x PBST at RT for 2 hr and incubated with primary antibody overnight at 4°C, followed by incubation with horseradish peroxidase-conjugated secondary antibodies for 1 hr at RT. Specific bands were visualized with enhanced chemiluminescence (ECL) substrate (ThermoFisher, #32106) and exposed onto films with an AX 700LE film processor (ALPHATEK).

DATA AND SOFTWARE AVAILABILITY

The RNA-seq datasets for gene expression in MCF-7 cells after siRNA-mediated knockdown of *EPIC1* (accession no. GSE98538), DNA methylation datasets of breast cancer cells (accession no. GSE57342 and GSE44837), and gene expression profile of breast cancer patients (accession no. GSE20711, GSE21653, GSE17907, GSE20685, GSE16446, and GSE19615), are available at GEO: https://www.ncbi.nlm.nih.gov/geo/.