THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 11, pp. 6857–6867, March 13, 2015

Long Non-coding RNAs (LncRNA) Regulated by Transforming Growth Factor (TGF) β

LncRNA-HIT-MEDIATED TGFβ-INDUCED EPITHELIAL TO MESENCHYMAL TRANSITION IN MAMMARY EPITHELIA[*S]

Received for publication, September 12, 2014, and in revised form, January 14, 2015 Published, JBC Papers in Press, January 20, 2015, DOI 10.1074/jbc.M114.610915

Edward J. Richards[‡], Gu Zhang[§], Zhu-Peng Li[§], Jennifer Permuth-Wey[¶], Sridevi Challa[‡], Yajuan Li[‡], William Kong[‡], Su Dan[§], Marilyn M. Bui[‖], Domenico Coppola[‖], Wei-Min Mao[§], Thomas A. Sellers[¶], and Jin Q. Cheng[‡1]

From the Departments of [‡]Molecular Oncology, [¶]Cancer Epidemiology, and [‖]Anatomic Pathology, Lee Moffitt Cancer Center and Research Institute, Tampa, Florida 33612 and the [§]Zhejiang Cancer Hospital & Zhejiang Cancer Research Institute, Zhejiang 310022, China

Background: Long noncoding RNAs (LncRNA) are emerging as key regulators in various biological processes. However, their role in epithelial-to-mesenchymal transition (EMT) remains elusive.
Results: A subset of lncRNAs are dysregulated upon transforming growth factor (TGF) β-induced EMT, and lncRNA-HIT mediates this process.
Conclusion: LncRNAs such as lncRNA-HIT ((HOXA transcript induced by TGFβ) play a pivotal role in EMT and breast cancer progression.
Significance: Here we profiled lncRNAs in TGFβ-induced EMT and identified a novel conserved lncRNA-HIT.

Long noncoding RNAs (lncRNAs) are emerging as key regulators in various biological processes. Epithelial-to-mesenchymal transition (EMT) is a developmental process hijacked by tumor cells to depart from the primary tumor site, invade surrounding tissue, and establish distant metastases. Transforming growth factor β (TGFβ) signaling has been shown to be a major inducer of EMT and to facilitate breast cancer metastasis. However, the role of lncRNAs in this process remains largely unknown. Here we report a genome-wide lncRNA profile in mouse mammary epithelial NMuMG cells upon TGFβ induction of EMT. Among 10,802 lncRNAs profiled, over 600 were up-regulated and down-regulated during the EMT, respectively. Furthermore, we identify that lncRNA-HIT (HOXA transcript induced by TGFβ) mediates TGFβ function, *i.e.* depletion of lncRNA-HIT inhibits TGFβ-induced migration, invasion, and EMT in NMuMG. LncRNA-HIT is also significantly elevated in the highly metastatic 4T1 cells. Knockdown of lncRNA-HIT in 4T1 results in decrease of cell migration, invasion, tumor growth, and metastasis. E-cadherin was identified as a major target of lncRNA-HIT. Moreover, lncRNA-HIT is conserved in humans and elevated expression associates with more invasive human primary breast carcinoma. Collectively, these data suggest that a subset of lncRNAs such as lncRNA-HIT play a significant role in regulation of EMT and breast cancer invasion and metastasis, and could be potential therapeutic targets in breast cancers.

Long noncoding RNAs (lncRNA)[2] are transcripts greater than 200 nucleotides that contain no open reading frame and lack protein coding capacity. Although they are much less conserved than protein coding genes and microRNA, accumulating evidence suggests that lncRNAs function in a broad range of cellular processes, such as cell growth, survival, migration, invasion, and differentiation (1–4). The TGFβ pathway is in part responsible for the epithelial to mesenchymal transition (EMT), a process by which primary epithelial cells acquire mesenchymal gene signatures to become more motile and invasive eventually leading to metastasis. LncRNAs regulated by TGFβ and their contribution to EMT has yet to be established in mammary epithelial cells.

TGFβ binds to a heteromeric complex of transmembrane serine/threonine kinases, the type I and II TGFβ receptors (TβRI and TβRII). Following ligand binding to TβRII, the type I receptor is recruited to the ligand-receptor complex, where the constitutively active TβRII transactivates TβRI. Activated TβRI phosphorylates the receptor-specific Smad2 and Smad3. Phosphorylated Smad2/Smad3 associates with Smad4 as a heteromeric complex and translocates to the nucleus, leading to the transcriptional induction or repression of a diverse array of genes (5). Previously, we and other have shown that protein coding genes and microRNAs that are regulated by the TGFβ pathway are functionally important in driving EMT and breast cancer metastasis (6). Therefore, we hypothesized that lncRNAs could similarly be regulated in response to TGFβ, and play a role in breast tumor progression.

In this report, we profiled changes of lncRNAs in NMuMG cells following TGFβ induction of EMT. AK020562, an uncharacterized lncRNA that locates in the Hoxa gene cluster, was significantly induced by TGFβ, and thus was named lncRNA-

* This work was supported, in whole or in part, by National Institutes of Health Grants CA137041, CA160455 (to J. Q. C.), and CA114343-03S1 (to T. A. S.) and the National Natural Science Foundation of China, Grant 81172081 (to W. M. M.).
[S] This article contains supplemental Tables S1–S3.
[1] To whom correspondence should be addressed: 12902 Magnolia Dr., SRB3, Tampa, FL 33612. Tel.: 813-745-6915; Fax: 813-745-3829; E-mail: jin.cheng@moffitt.org.

[2] The abbreviations used are: lncRNA, long noncoding RNA; EMT, epithelial to mesenchymal transition; HOXA, homeobox domain A; HIT, Hoxa-induced TGFβ; LNAS, locked nucleic acid; qPCR, quantitative PCR.

HIT. Depletion of lncRNA-HIT inhibits TGFβ-induced migration, invasion, and EMT in NMuMG, and decreases primary tumor growth and metastasis in a 4T1 orthotopic mouse xenograft model. In contrast to the protein coding genome, the conservation of lncRNA across species remains poorly defined and not very well understood (7, 8). Despite this discrepancy, lncRNA-HIT is well conserved in sequence and chromosomal location from mouse to human. We were able to detect lncRNA-HIT in human and observed that increased expression directly correlates with breast cancer progression.

EXPERIMENTAL PROCEDURES

Cell Line and Treatment—NMuMG epithelial cells were purchased from the American Type Culture Collection (Manassas, VA). 4T1, 4T07, 168FARN, and 67NR cells were a kind gift of Fred Miller (Wayne State University). All cells were grown in complete medium, DMEM (Invitrogen) containing 10% fetal bovine serum (FBS) supplemented with 1 mM L-glutamine, penicillin/streptomycin, and non-essential amino acids (Gibco). NMuMG cells were treated with TGFβ at a concentration of 5 ng/ml for the times indicated in the figures and legends. Cell transfection experiments were performed with Lipofectamine 2000 (Invitrogen).

LncRNA and mRNA Microarray—NCode™ Mouse Noncoding Microarray chip from Life Technologies (Carlsbad, CA), which contains 10,802 lncRNAs and 25,178 protein-coding genes, was used to interrogate lncRNA and mRNA changes in vehicle- *versus* TGFβ (5 ng/ml)-treated NMuMG cells after 24 h. TFGβ was purchased from R&D Systems (Minneapolis, MN). Total RNA was isolated by TRIzol (Life Technologies), end labeled, and hybridized to array. Hybridization and analysis were performed in the Molecular Genomics Core at the H. Lee Moffitt Cancer Center.

RNA Isolation and Reverse Transcription-Quantitative Polymerase Chain Reaction (RT-qPCR)—Total RNAs were isolated with TRIzol reagent following the manufacturer's protocol and then subjected to RT reaction using a High Capacity cDNA Reverse Transcription Kit (Life Technologies). The RT product was used for subsequent qPCR. The qPCR was performed with SYBR Green 2× Master Mix (Life Technologies) on ABI HT9600 from Applied Biosystems (Foster City, CA) and data were collected and analyzed using ABI SDS version 2.3. ΔC_T values were normalized to GAPDH, and $\Delta\Delta C_T$ analysis was performed to calculate relative RNA expression. RT-qPCR primers and siRNAs were obtained from Integrated DNA Technologies (Coralville, IA) and their oligonucleotide sequences for array validation are listed in supplemental Table S3. LncRNA-HIT RT-qPCR primer sequences are as follows: forward 5′-GAATTGGTCTCAGCCAGATTCG-3′ and reverse 5′-GTCAAGTCAGAGCTCAGTTCCC-3′. siRNAs sequences designed to target lncRNA-HIT are as follows: si-HIT-1, antisense 5′-CUAGAGUUAGGAAGGUAUGUGAGACUU-3′ and sense, 5′-GUCUCACAUACCUUCCUAACUCUAG-3′ and si-HIT-2, antisense 5′-UUAAGGUCACAGACCACCUUGGAGGGU-3′ and sense, 5′-CCUCCAAGGUGGUCUGUGACCUUAA-3′.

Plasmids—Full-length lncRNA-HIT cDNA was amplified using the FastStart High Fidelity PCR System (Roche Applied Science) and NMuMG genomic DNA as a template. Cloning primers contained EcoRI sites for inserting cDNA into the pcDNA3.1(+) expression vector. Primer sequences used to amplify full-length lncRNA-HIT are as follows: forward, 5′-GGAATTCAGGTTTCGAGGAGATGAAAGTG-3′ and reverse, 5′-GGAATTCTCTGAACACAATATCTCCTGTGC-3′. Plasmid was sequenced validated for PCR introduced errors and insert orientation following selection of positive clone. hE-cadherin-pcDNA3 was a kind gift from Barry Gumbiner and purchased through Addgene (9).

Invasion and Migration Assays—NMuMG and 4T1 cell lines were transfected with control siRNA (si-CTL), si-HIT-1 or -2. After transfection for 48 h, NMuMG and 4T1 cells were treated with TGFβ (5 ng/ml) or vehicle control for 24 h and then seeded into the upper chamber of Boyden Chambers coated without (migration) and with (invasion) Matrigel. Top chambers contained serum-free media, whereas lower chambers had media containing 10% fetal bovine serum. After 16 h, invasion and migration were evaluated and quantified by estimating the mean ± S.D. of 4 non-biased image fields.

Immunofluorescence, Immunoblotting, and Antibodies—The cells were grown to 60% confluence on coverslips and transfected with si-CTL or si-HIT-1 or -2 for 48 h. Subsequently, TGFβ (5 ng/ml) was added to siRNA-transfected cells for the range of times indicated in the figure legends. Briefly, cells were washed with PBS, fixed with 10% formalin containing methanol, and permeabilized with 1% Nonidet P-40 in PBS. Cells were blocked in 10% normal goat serum for 1 h, and 1:200 dilution of primary antibodies were incubated at 4 °C overnight. Coverslips were washed in PBS 3 times and then appropriate secondary antibodies were added at 1:500 dilution. Coverslips were washed in TBS 3 times, counterstained with DAPI, and fixed for visualization. Western blot was performed as previously described (10). The band intensity of Western blots was quantified using ImageJ software and represented as ratio of target gene/β-actin. Antibody against E-cadherin was purchased from BD Transduction Labs (San Jose, CA), Lamin A/C from Santa Cruz Biotechnology (Santa Cruz, CA), and antibodies for Vimentin and GAPDH from Cell Signaling Technologies (Danvers, MA). Alexa Fluor secondary anti-mouse and anti-rabbit antibodies for immunofluorescence were purchased from Life Technologies, and HRP-linked antibody for Western blot analysis was purchased from Sigma.

Luciferase Reporter Assay—E-cadherin promoter luciferase plasmid was co-transfected with vector or pcDNA-HIT for 48 h in NMuMG. Cells were then washed in PBS, lysed, and then firefly substrate was added to detect expression of luciferase activity. The reaction was stopped using *Renilla* substrate as a control and performed using the Promega Dual Luciferase Assay Kit (Madison, WI).

Soft Agar Colony Formation Assay—4T1 cells were transfected with si-CTL, si-HIT-1 or -2 for 48 h then washed, trypsinized, and resuspended in RPMI. A bottom layer of 0.6% agar in RPMI was poured and a layer of 0.3% agar in RPMI was layered on top containing 1,000,000 per 12-well and allowed to grow for 2 weeks. Four replicates were plated for each condition and quantitation is represented as mean ± S.D, of colonies counted in 4 non-biased fields.

Orthotopic Xenograft Model—4T1 cells were transfected with si-CTL and si-HIT-1 for 48 h then washed, trypsinized, and resuspended in PBS at a concentration of 4×10^6 cells/100

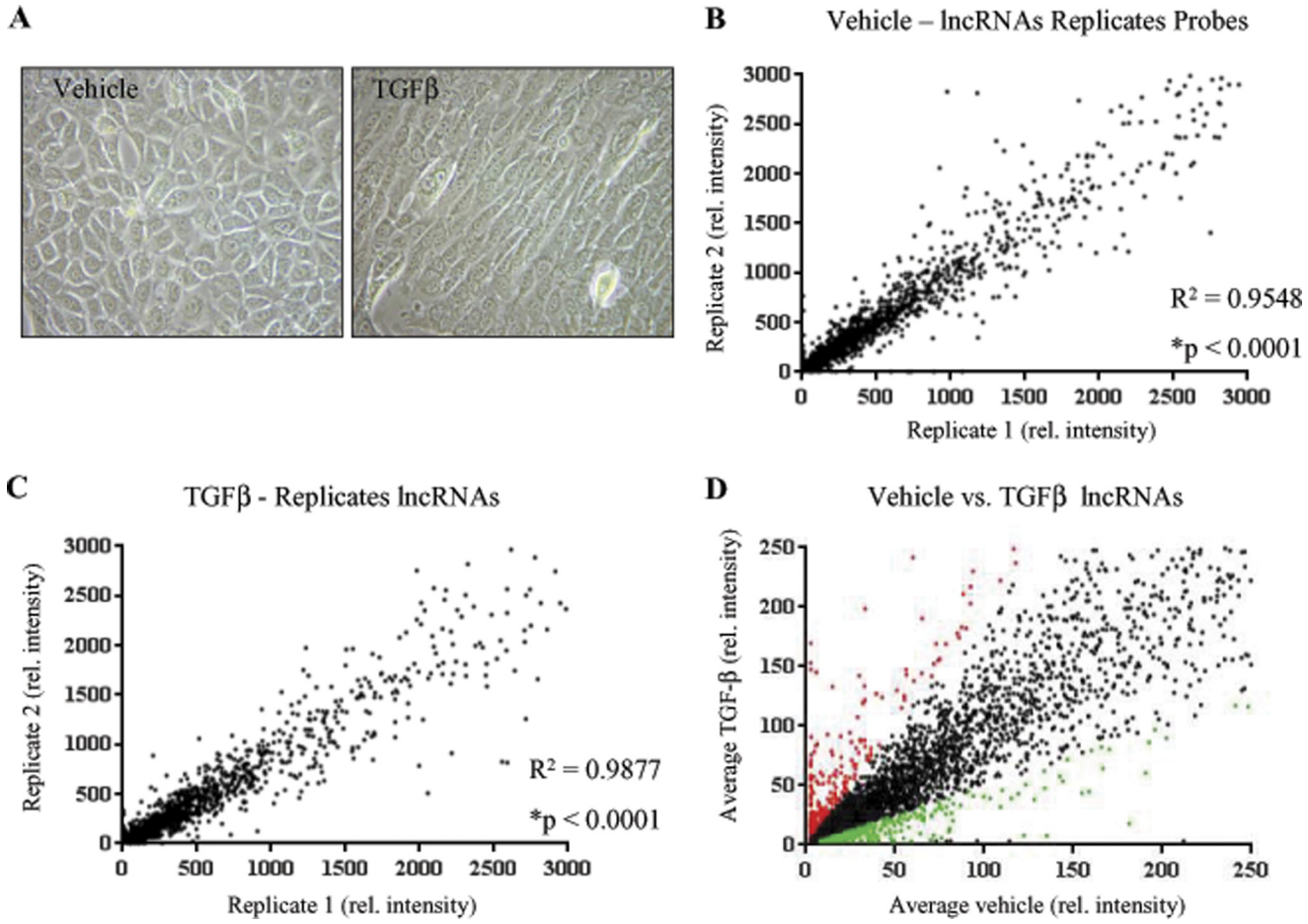


FIGURE 1. **LncRNA expression profile of TGFβ-induced EMT in NMuMG cells.** *A*, TGFβ induces a cell morphological change of EMT in NMuMG cells. The cells were treated with vehicle (*left*) or 5 ng/ml of TGFβ (*right*) for 24 h and then photographed. *B–D,* linear regression analysis of lncRNAs in 2 replicates after vehicle (*B*) and TGFβ (*C*) treatment of NMuMG cells for 24 h. *Panel D* is a scatter plot analysis comparing all lncRNA changes of vehicle *versus* TGFβ-treated NMuMG (*D*). Note: *red* and *green dots* represent up-regulated and down-regulated lncRNAs, respectively.

μl. Cells were mixed with a 1:1 ratio with Matrigel from Corning (Manassas, VA) and injected into the lower mammary fat pad of Nu/Nu mice from Charles River (Wilmington, MA). Primary tumor growth was monitored via standard caliper measurements. At 4 weeks end point mice primary tumors were weighed and the lungs were removed, inflated with 10% buffered formalin, and stained using Bouin solution.

Locked Nucleic Acid in Situ Hybridization of Formalin-fixed, Paraffin-embedded Tissue Microarray—LncRNA-HIT locked nucleic acid (LNA) probe was prepared by 5′ end labeling with digoxigenin-ddUTP terminal transferase using the DIG 5′ End Labeling Kit from Roche Applied Science. Probe sequence containing LNAs (+) is as follows: 5′-AATGGCA+G+A+T+T+C+A+C+AAGCATCA-3′. Following deparaffinization and proteinase K digestion, breast tumor tissue microarrays were prehybridized for 1 h and then hybridized with 10 nmol/liter of LNA lncRNA-HIT probe in a hybridization buffer (Roche Applied Science) for 12 h. After three consecutive washes in 4× SSC, 50% formamide, 2× SSC, and 0.1 × SSC, sections were treated with a blocking buffer (Roche Applied Science) for 1 h and incubated with anti-DIG-AP Fab fragments (Roche Applied Science) for 12 h. After washing three times in 1× maleic acid and 0.3% Tween 20 buffer, reactions were processed in a detection solution (100 mmol/liter of Tris-HCl (pH 9.5) and 100 mmol/liter of NaCl) in the presence of nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate from Promega and then visualized under a microscope. Allred scoring system was used for quantification (11).

Statistical Analysis—Statistical significance was determined using unpaired Student's *t* test, and $p \leq 0.05$ was considered to be statistically significant.

RESULTS

Profile of lncRNA Expression in TGFβ-induced EMT in NMuMG Cells—Using the widely employed NMuMG mammary epithelial cells as a model to study TGFβ-induced EMT, we have previously shown that TGFβ up-regulates ncRNA BIC, which processes into miR-155 to contribute to EMT (6). Therefore, we hypothesized that TGFβ could similarly regulate lncRNAs that control EMT-associated phenotypes. To address this, we treated NMuMG cells with vehicle or TGFβ (5 ng/ml) for 24 h to induce EMT (Fig. 1*A*). Total RNA was labeled with digoxigenin and hybridized to NCode Mouse Noncoding RNA Microarray. This is a high-density oligoarray that contains lncRNAs and protein coding genes to detect the dynamic expression of a large subset of lncRNAs and associated protein coding genes within the same biological sample. This array contains 2 replicates for each gene per sample to ensure signal

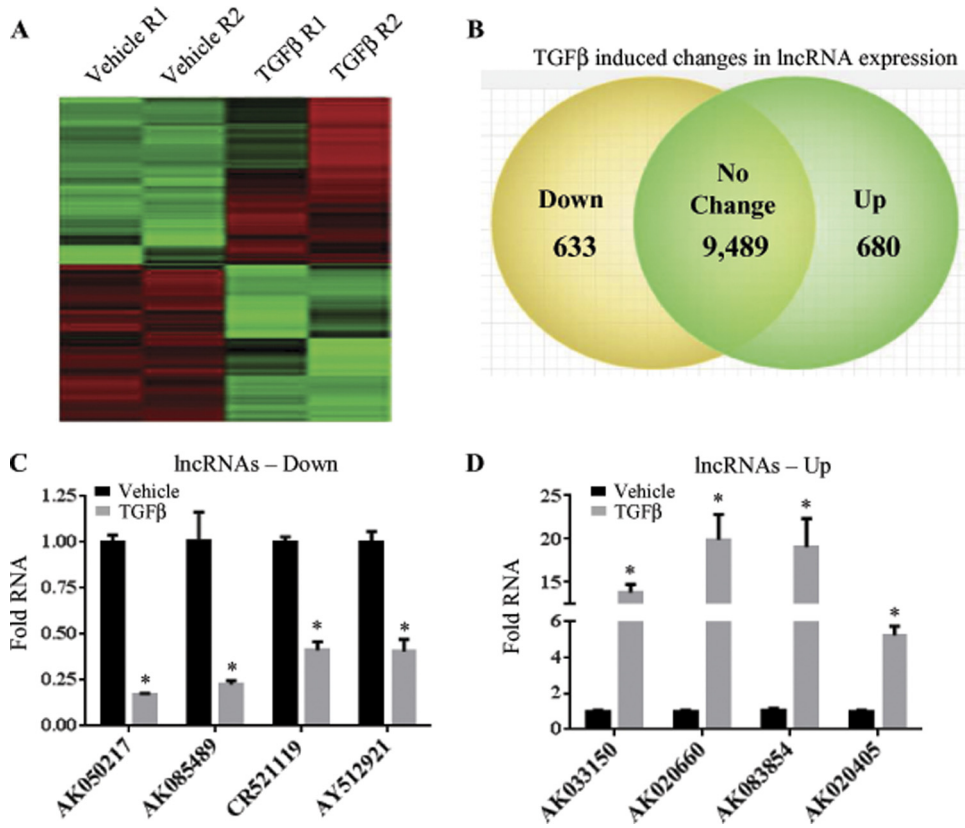


FIGURE 2. **Verification of TGFβ-regulated lncRNAs.** Heat map (*A*) and diagram (*B*) representations of dysregulated lncRNAs between vehicle- and TGFβ-treated NMuMG cells. Real-time PCR analysis of representative TGFβ up-regulated (*C*) and down-regulated (*D*) lncRNAs.

TABLE 1
List of top 15 lncRNAs down-regulated by TGFβ

Gene ID	Vehicle (average relative intensity)	TGFβ (average relative intensity)	Average fold-change	Chromosome	Length
					nt
AK032278	38,274.10	3.45	0.0001	1	4,151
AK082403	380.93	3.45	0.009	5	1,925
AK076494	251.28	3.45	0.013	18	2,377
AK016398	211.75	3.45	0.016	14	1,401
AK083030	313.03	6.04	0.019	10	1,958
AK084380	113.95	3.45	0.030	16	4,202
AK036704	95.84	3.55	0.037	10	2,401
AK087232	78.73	3.45	0.043	4	3,346
AK020718	80.89	3.59	0.044	X	644
AK080421	73.04	3.45	0.047	2	1,108
AK039030	69.51	3.99	0.057	1	3,090
AK019328	134.91	7.95	0.059	16	205
AK007385	58.94	3.50	0.059	3	839
AK141015	70.40	4.22	0.059	X	2,752
AK015950	122.10	8.16	0.066	1	483

detection is consistent and reliable. Plots of relative intensity signals of replicate spots for lncRNAs in vehicle and TGFβ treatment indicate the consistency of signal for lncRNA within given sample. Linear regression analysis provides statistically significant evidence that both vehicle and TGFβ-treated replicates are consistent (Fig. 1, *B* and *C*). The same analysis was also performed and demonstrated to be significant for protein coding genes (data not shown).

The microarray contains replicates for 10,802 lncRNAs in which the signal was normalized, averaged, and plotted to assess overall changes in expression in vehicle *versus* TGFβ treatment (Fig. 1*D*). We determined that lncRNA expression was significantly changed if the average signal of replicates changed ≥2-fold from vehicle to TGFβ-treated samples (*red* and *green spots* in Fig. 1*D*). Data including average signal and fold-change for all lncRNAs are summarized (supplemental Table S1). Microarray analysis was also performed for 25,178 protein coding genes and summarized in supplemental Table S2). Using ≥2-fold change as the cutpoint, 633 lncRNAs were down-regulated and 680 lncRNAs were up-regulated following TGFβ-induced EMT (Fig. 2, *A* and *B*). Tables 1 and 2 summarize the top 15 down-regulated and 15 up-regulated lncRNAs. Furthermore, we confirmed the expression of several top deregulated lncRNAs by real-time PCR (Fig. 2, *C* and *D*). In total, 1,096 genes increased and 2,220 decreased expression upon TGFβ-induced EMT (Fig. 3, *A* and *B*). Furthermore, RT-qPCR validation

TABLE 2
List of top 15 lncRNAs up-regulated by TGFβ

Gene ID	Vehicle (average relative intensity)	TGFβ (average relative intensity)	Average fold-change	Chromosome	Length
					nt
AK171786	3.19	23,171.79	7257.60	18	3,372
AK043331	8.67	8287.23	954.83	9	2,462
AK160425	2.48	169.80	68.27	5	2,013
AK085133	2.28	153.48	67.25	3	2,662
AK014868	2.46	147.81	60.06	X	1,979
AK039984	2.67	90.99	34.03	11	986
AK020562	2.42	68.74	28.38	6	1,245
AK005664	2.61	71.50	27.32	2	652
AK143298	2.49	66.65	26.69	12	3,389
AK035982	5.68	145.23	25.56	5	1,388
AK015200	2.39	58.93	24.59	5	737
AK039295	2.45	55.29	22.53	6	1,385
AK020168	2.67	57.04	21.34	3	1,196
AK019733	2.62	55.21	21.00	10	355
AK016032	2.46	50.17	20.32	7	1,118

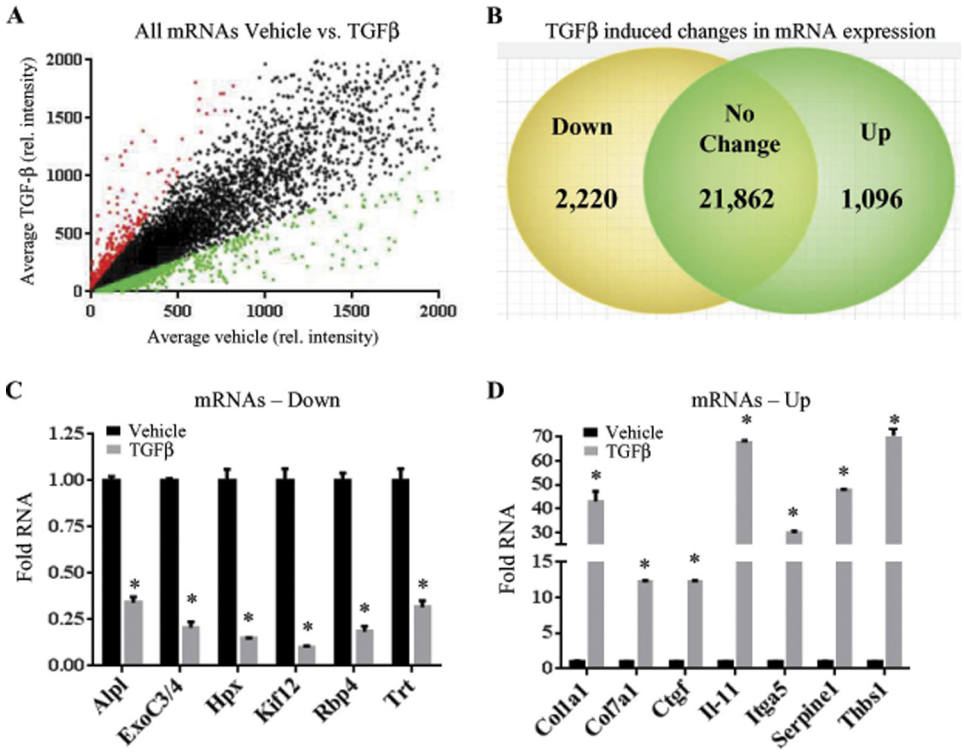


FIGURE 3. **Dysregulated mRNAs by TGFβ.** *A,* scatter plot analysis comparing all mRNA changes of vehicle *versus* TGFβ-treated NMuMG cells (*red* and *green dots* represent up-regulated and down-regulated mRNAs, respectively). *B,* diagram summarizing the number genes changed. *C* and *D*, real-time PCR analysis of representative TGFβ down-regulated (*C*) and up-regulated (*D*) mRNAs. *Asterisks* represent $p < 0.05$.

showed significant changes of several established typical TGFβ target protein-coding genes including *Col1a1*, *Ctgf*, *Itga5*, *Serpine1*, *Kif12*, *Rbp4*, etc. (Fig. 3, *C* and *D*), indicating that the lncRNA data generated in this study are dependable.

LncRNA-HIT Mediates TGFβ-induced Invasion, Migration, and EMT in NMuMG Cells—LncRNA-HIT (AK020562) is one of the top TGFβ up-regulated lncRNAs, and was of particular interest to us given its genomic location, *i.e.* within the homeobox domain A (Hoxa) gene cluster. LncRNA-HIT resides in the sense orientation to the Hoxa protein genes and downstream of Hoxa13. There is no overlap between these transcripts and a gap of ~1.4 kb between the Hoxa13 3′ termini and lncRNA-HIT 5′ transcriptional start site (Fig. 4*A*). Previous studies have shown that the HOX gene clusters (HOXA, HOXB, HOXC, and HOXD) are well conserved across species and several lncRNAs have been shown to be functionally important within these regions (1, 12, 13).

Consistent with the microarray finding, RT-qPCR analysis showed that lncRNA-HIT expression was induced by TGFβ (Fig. 4*B*). LncRNAs have been shown to play functional roles in both the nuclear and cytoplasmic compartments (14–17). Therefore, we performed cellular fractionation after 24 h of treatment with TGFβ to assess the subcellular localization of lncRNA-HIT. We observed that lncRNA-HIT expression was induced and remained localized to the nucleus upon TGFβ treatment (Fig. 4*C*), suggesting that lncRNA-HIT functions in the nucleus. We further designed 2 siRNAs against lncRNA-HIT (si-HIT-1 and si-HIT-2) to experimentally address its functional significance in TGFβ-induced migration, invasion, and EMT. NMuMG cells were transfected with scrambled control siRNA (si-CTL), si-HIT-1, or si-

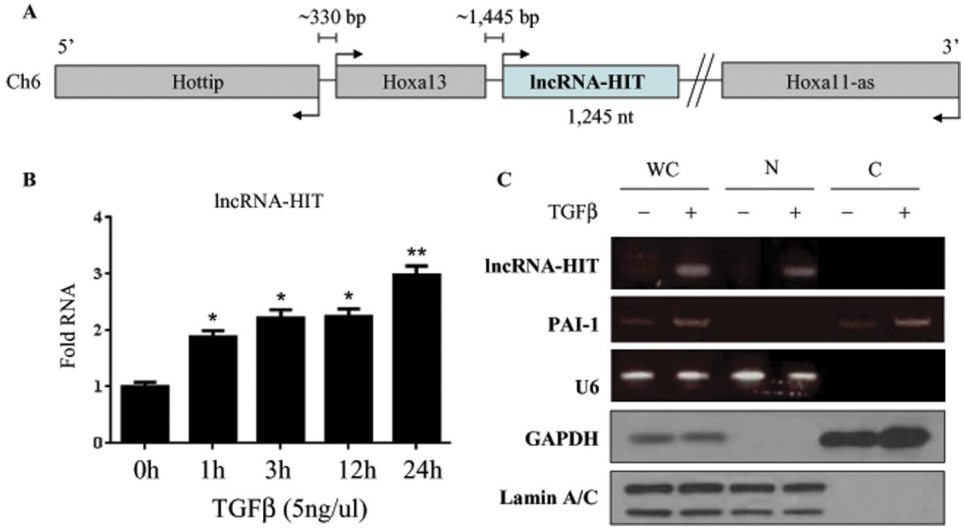


FIGURE 4. **LncRNA-HIT locates in Hoxa gene cluster and is induced by TGFβ.** *A,* diagram shows lncRNA-HIT location. *B*, RT-qPCR revealed that lncRNA-HIT is progressively induced in NMuMG cells after tracking expression over a 24-h time course of TGFβ treatment. *C,* lncRNA-HIT is predominantly localized in the nucleus. Following treatment with vehicle or TGFβ for 24 h, NMuMG cells were fractionated and then subjected to semiquantitative PCR (*top*), including immunoblot analysis with the indicated control antibodies and semi-quantitative PCR with control primers (medium and bottom). *Asterisks* represent $p < 0.05$.

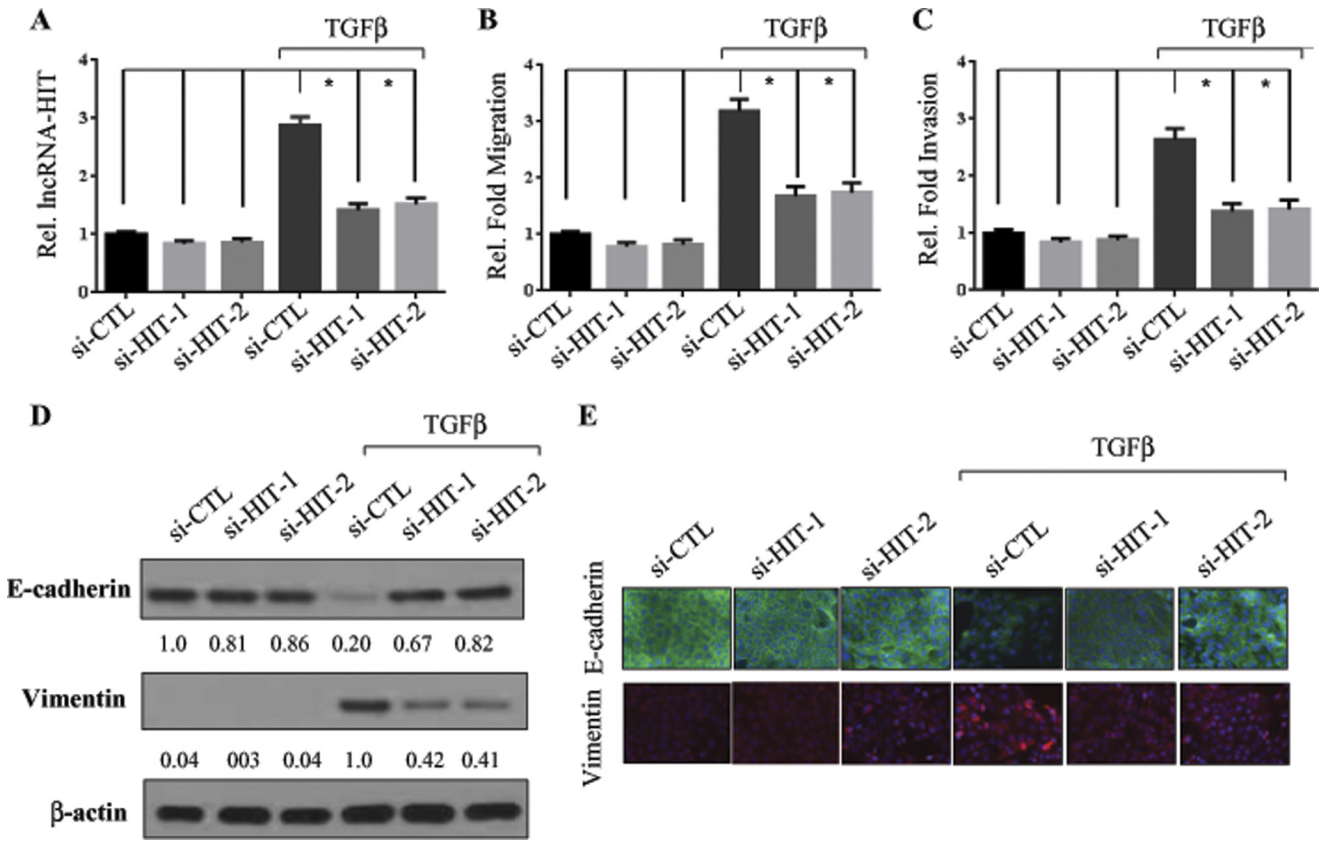


FIGURE 5. **LncRNA-HIT mediates TGFβ-induced cell migration, invasion, and EMT.** *A*, NMuMG cells were transfected with indicated siRNAs. After treatment with vehicle or TGFβ for 48 h, RNAs were isolated and analyzed for lncRNA-HIT expression using real-time PCR. *B* and *C,* depletion of lncRNA-HIT reduces TGFβ-triggered cell migration and invasion. LncRNA-HIT knockdown and control NMuMG cells were seeded in a Boyden chamber without (*B*) and with (*C*) Matrigel to assess changes in migration and invasion following the procedure described under "Experimental Procedures." *D* and *E,* knockdown of lncRNA-HIT abrogates a hallmark change of EMT induced by TGFβ. NMuMG cells were transfected with the indicated siRNAs. After 48 h of transfection, the cells were left untreated or treated with TGFβ for 16 h. Expression of E-cadherin and Vimentin were assessed by Western blot (*D*). Immunofluorescence staining was also performed with antibodies against E-cadherin and Vimentin (*E*). Note: TGFβ-induced tight junction dissolution was largely rescued by depletion of lncRNA-HIT. *Asterisks* represent $p < 0.05$.

HIT-2 for 48 h. Cells were then treated with TGFβ (5 ng/ml) or vehicle control for 24 h. Following confirmation of knockdown of lncRNA-HIT (Fig. 5*A*), we performed two-chamber migration and invasion assays and found that depletion of lncRNA-HIT significantly inhibited TGFβ-induced migration and invasion compared with si-CTL (Fig. 5, *B* and *C*).

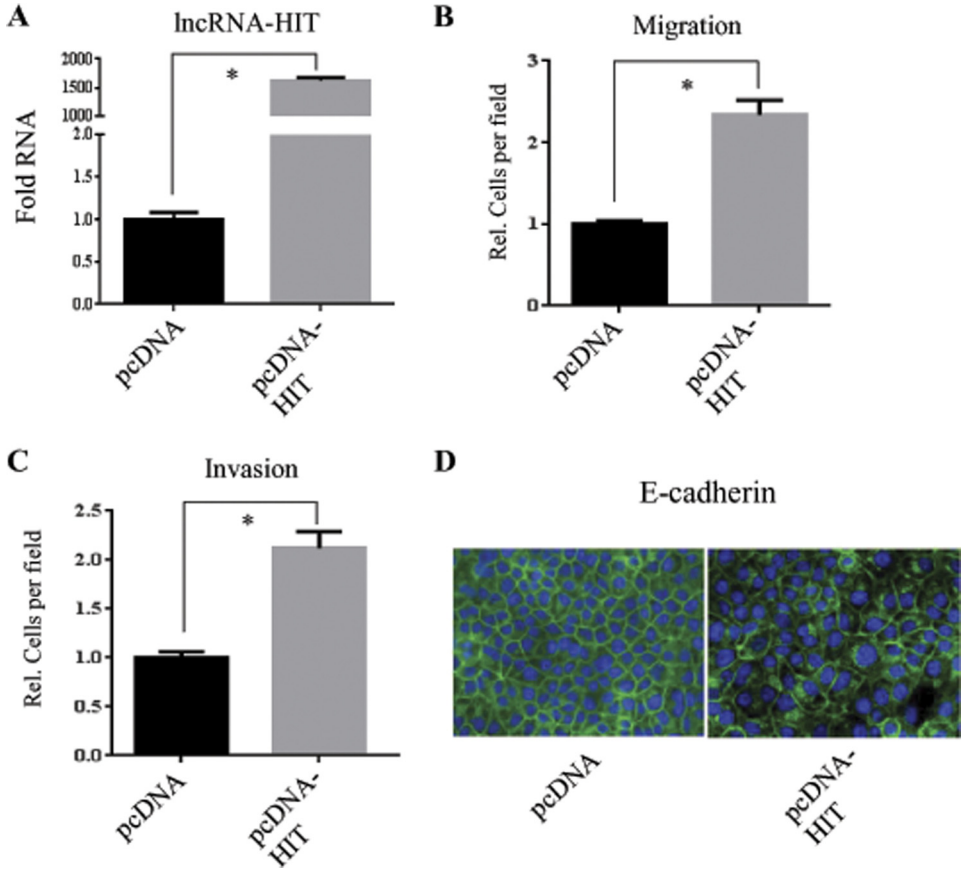


FIGURE 6. **Overexpression of lncRNA-HIT promotes migration, invasion and disrupts tight junction.** *A,* lncRNA-HIT was ectopically expressed in NMuMG cells for 48 h. Cells were then seeded in the upper Boyden chamber without (*B*) and with (*C*) Matrigel to determine changes in migration and invasion. *D,* presence of tight junction following overexpression of lncRNA-HIT was assessed using E-cadherin antibody and performing immunofluorescence staining.

TGFβ-induced EMT in NMuMG causes the cell to move from an epithelial to a more mesenychmal and fibroblast-like gene signature. A hallmark of EMT is the loss of E-cadherin and increase of Vimentin expression. To test whether lncRNA-HIT can promote the process of EMT, we transfected the cells for 48 h with si-CTL, si-HIT-1, and si-HIT-2 and then treated cells with and without TGF-β for 24 h to induce EMT. Immunoblotting analysis revealed that TGFβ treatment reduced E-cadherin and increased Vimentin in si-CTL cells. However, knockdown of lncRNA-HIT largely overrode the TGFβ action in E-cadherin and Vimentin expression (Fig. 5*D*). This result was recapitulated using immunofluorescence staining to visualize expression of E-cadherin and Vimentin (Fig. 5*E*). We noted that TGFβ-disrupted tight junctions were largely restored by knockdown of lncRNA-HIT (Fig. 5*E*). Moreover, we examined the effects of overexpression of lncRNA-HIT alone on EMT in NMuMG cells. In NMuMG cells overexpressing lncRNA-HIT (Fig. 6*A*), we observed a significant increase in both migration and invasion (Fig. 6, *B* and *C*). Ectopic expression of lncRNA-HIT was also able to disrupt tight junction as indicated by E-cadherin immunofluorescence staining (Fig. 6*D*). Collectively, these data indicate that lncRNA-HIT plays a pivotal role in TGFβ-induced cell migration, invasion, and EMT.

LncRNA-HIT Is Up-regulated in 4T1 Cells and Its Depletion Inhibits Cell Migration, Invasion, Lung Metastasis, and Tumor Growth—We further examined lncRNA-HIT expression in four well characterized mouse mammary tumor cell lines (67NR, 168FARN, 4TO7, and 4T1) derived from a single spontaneously arising mammary tumor in a BALB/c mouse. Although each of these tumor cell lines is able to form primary tumors, they have different metastatic properties. 67NR cells form primary tumors readily, but tumor cells do not intravasate. 168FARN cells can be detected in lymph nodes but rarely in other tissues, suggesting that they can enter the vasculature, but extravasate inefficiently. 4TO7 cells can disseminate from primary mammary tumors into the lungs but do not form visible lung nodules. Moreover, disseminated 4TO7 cells in the lungs rapidly disappear when the primary tumor is removed, suggesting that they are unable to colonize distant sites. 4T1 cells are fully metastatic and form macroscopic lung nodules from primary mammary tumors. Interestingly, lncRNA-HIT was significantly elevated in the 4T1 cell line and expressed at low levels in other 3 cell lines (Fig. 7*A*), suggesting that lncRNA-HIT is involved in breast cancer metastasis. The importance of lncRNA-HIT in 4T1 cell migration and invasion was further investigated. After cells were transfected with si-HIT-1, and si-HIT-2 as well as si-CTL for 48 h (Fig. 7*B*), two-chamber migration and invasion assays were performed as described above. After 16 h, we observed a significant reduction in both migration and invasion of 4T1 cells in which lncRNA-HIT was depleted (Fig. 7, *C* and *D*). However, we did not observe that TGFβ-induced lncRNA-HIT in 4T1 cells and that knockdown of lncRNA-HIT had no effect on TGFβ-induced 4T1 cell migration and invasion (data not shown).

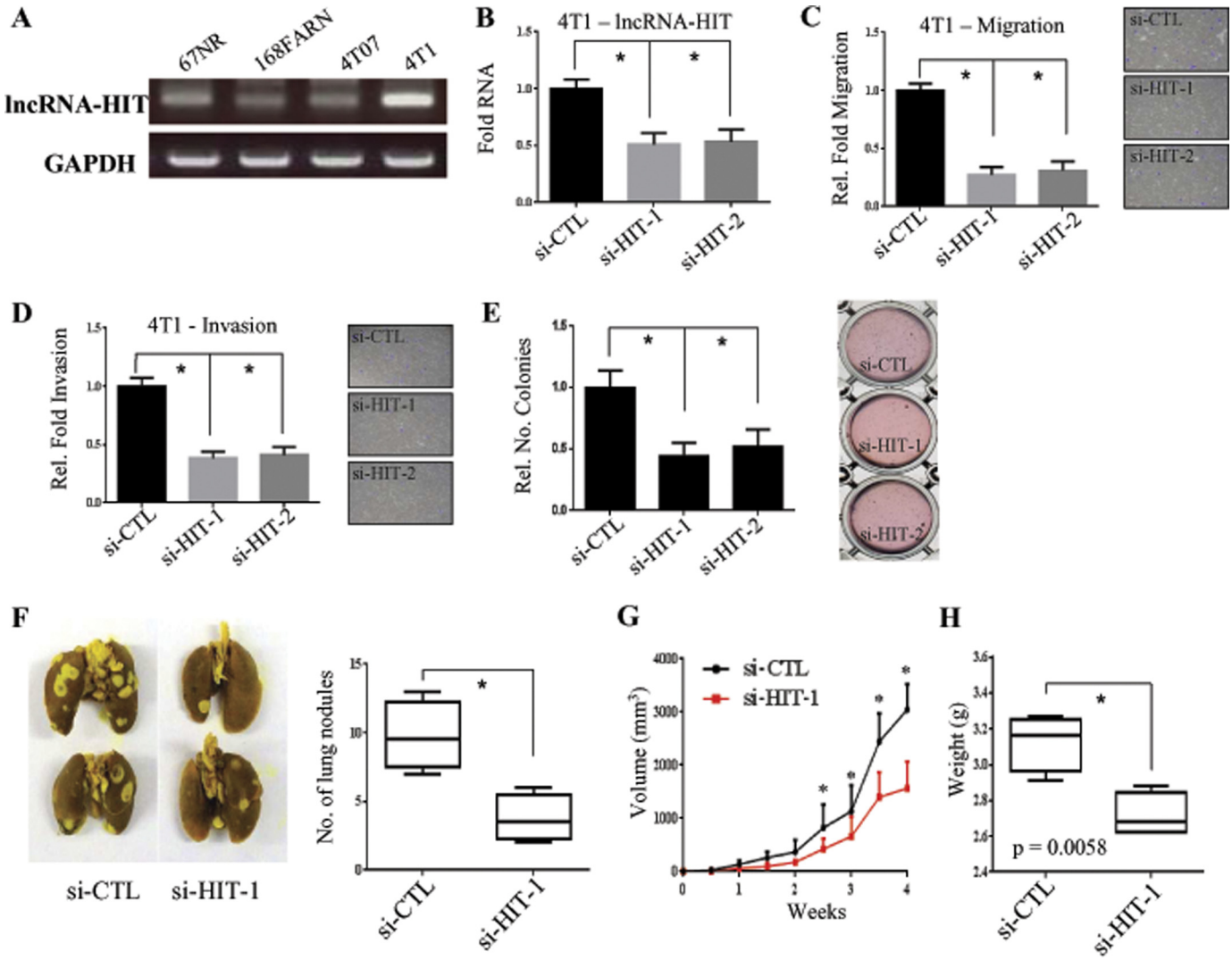


FIGURE 7. **LncRNA-HIT is elevated in highly metastatic 4T1 cells and its knockdown results in significant reduction of cell migration and invasion.** *A*, expression of lncRNA-HIT was evaluated in 4 isogenic mouse breast cancer cell lines 67NR, 168FARN, 4T07, and 4T1 by semiquantitative RT-PCR. *B*, knockdown of lncRNA-HIT with 2 siRNAs in 4T1 cells. Cells were transfected with the indicated siRNAs and then analyzed lncRNA-HIT expression by real-time PCR. *C* and *D*, depletion of lncRNA-HIT dramatically reduces 4T1 cell migration and invasion. Following treatment with the indicated siRNAs, 4T1 cells were assayed for cell migration and invasion as described in the legend to Fig. 4. *E*, soft agar colony formation assay was performed after 4T1 cells were transfected with si-CTL, si-HIT-1, or si-HIT-2. Colony forming capacity was dramatically decreased after 2 weeks in lncRNA-HIT-depleted cells. *F*, orthotopic breast cancer model. Control and lncRNA-HIT knockdown 4T1 cells (4×10^6) were injected to mammary fat pad of female nude mice and tumor growth was monitored for 4 weeks. The number of metastatic lung nodules (*F*), tumor volume (*G*), and tumor weight (*H*) were assessed at completion of experiment. *Asterisks* represent $p < 0.05$.

We next examined the effect of lncRNA-HIT on anchorage independent growth in soft agar following depletion of lncRNA-HIT. Compared with si-CTL-treated 4T1 cells, si-HIT-1 and -2 knockdown cells resuspended in soft agar and cultured for 2 weeks showed markedly reduced colony formation capacity (Fig. 7*E*). Notably, the orthotopic breast cancer model revealed that knockdown of lncRNA-HIT dramatically inhibited the number of metastatic lung nodules, breast tumor volume, and tumor weight (Fig. 6, *F*–*H*).

Identification of E-cadherin as a Major Target of lncRNA-HIT—To identify the genes regulated by lncRNA-HIT, we performed Affymetrix gene expression analysis after 48 h ectopic expression of lncRNA-HIT in NMuMG (Fig. 8*A*). One of the significant deregulated genes associated with EMT was E-cadherin. Accumulating studies have shown that loss of E-cadherin is not only a hallmark of EMT but also a key driver of EMT and metastasis (18–21). Thus, we further examined if lncRNA-HIT inhibits E-cadherin transcription by performing RT-qPCR and using the E-cadherin luciferase promoter assay. Following ectopic expression of lncRNA-HIT, we observed a significant loss of E-cadherin mRNA and promoter activity (Fig. 8, *B* and *C*). Furthermore, this correlated with a loss of E-cadherin protein after 72 h overexpression of lncRNA-HIT (Fig. 8*D*), however, we did not see changes of the other EMT-associated genes ZEB1 or Snail at this time point. Furthermore, the effects of lncRNA-HIT-induced EMT, migration, and invasion were rescued through introduction of ectopic E-cadherin (Fig. 8, *E* and *F*).

Expression of LncRNA-HIT Is Conserved and Is Elevated in Human Invasive Ductal Breast Cancer—LncRNA-HIT resides in the 5′ distal Hoxa gene cluster, a region of the genome that is highly conserved (12, 22). We therefore used BLAST analysis software to evaluate the conservation of lncRNA-HIT in

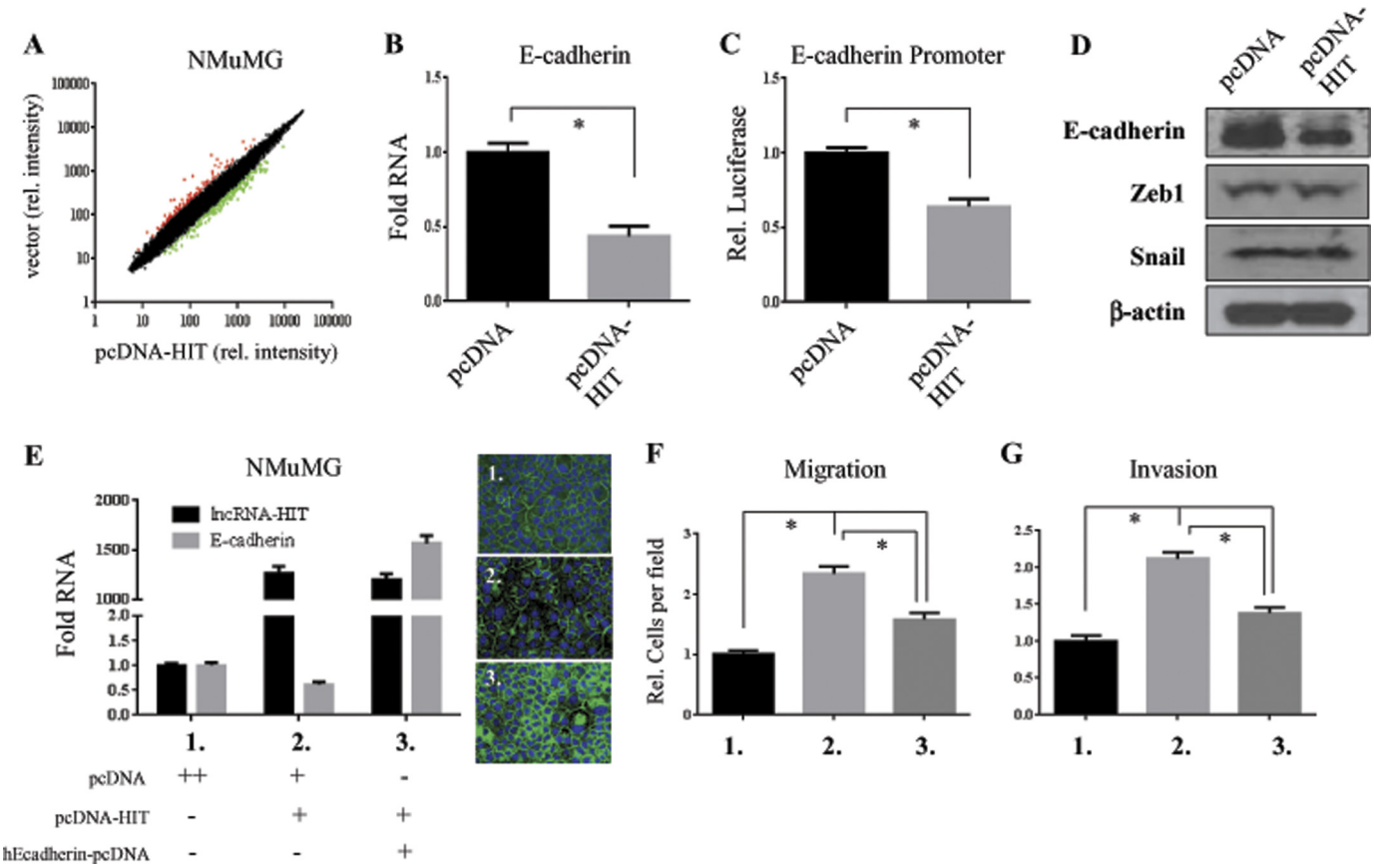
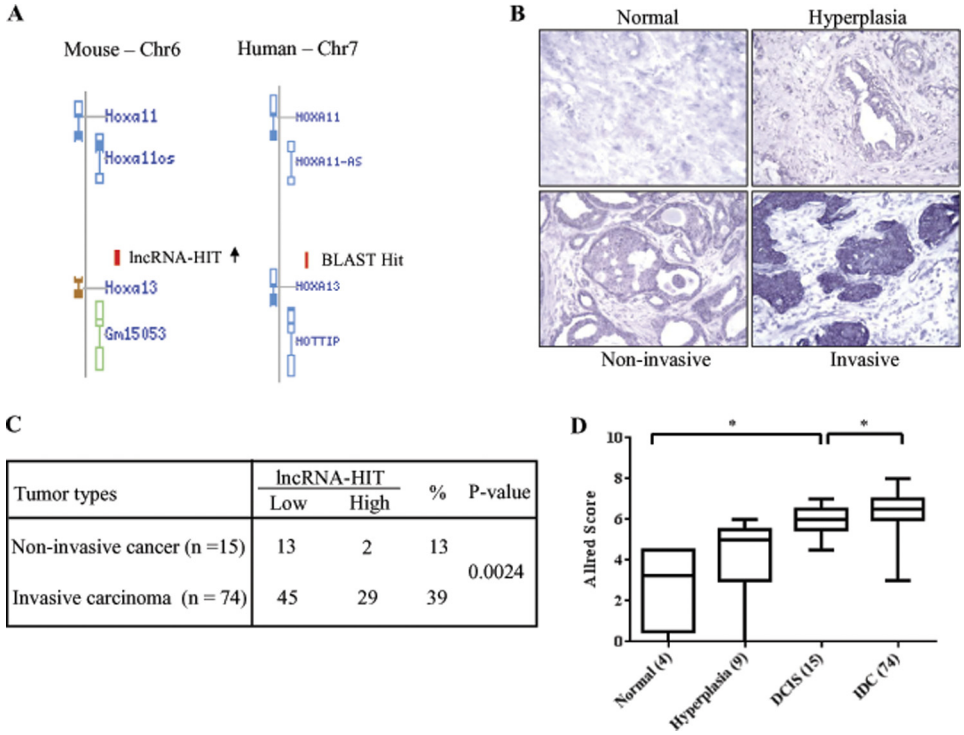


FIGURE 8. **E-cadherin is a target of lncRNA-HIT.** *A*, Affymetrix Gene Expression Analysis was performed on NMuMG cells transfected with vector and lncRNA-HIT. *B*, RT-qPCR was performed to confirm a loss of E-cadherin expression as indicated as down-regulated in the array. *C*, PGL3-E-cadherin promoter was co-transfected with vector or lncRNA-HIT into NMuMG cells for 48 h and then assayed for luciferase activity. *D*, after 72 h overexpression of lncRNA-HIT E-cadherin, as well as other EMT markers, ZEB1 and Snail, were immunoblotted for protein expression. *E*, after 72 h transfection, hE-cadherin-pcDNA expression was able to rescue loss of tight junction as indicated by E-cadherin immunofluorescence staining and lncRNA-HIT induced migration (*F*) and invasion (*G*). *Asterisks* represent $p < 0.05$.

humans and the potential for orthologs residing in the same genomic location. LncRNA-HIT maps to the syntenic loci (Fig. 9*A*) and shares 99% coverage and significant sequence identity (80%) to human. Having demonstrated that lncRNA-HIT mediates TGFβ-induced EMT in NMuMG and cell migration and invasion in 4T1 cells, we asked if the expression of lncRNA-HIT was associated with cancer invasiveness in primary breast carcinoma in human. A total of 89 breast cancer specimens (15 noninvasive and 74 invasive breast carcinomas) and 4 normal and 9 hyperplasia breast tissue samples were examined for the expression of lncRNA-HIT. Locked nucleic acid *in situ* hybridization (LNA-ISH) (Fig. 9*B*) analyses revealed high levels of lncRNA-HIT in 29 of 74 invasive tumors but in only 2 of 15 noninvasive cancer tissues (Fig. 9*C*). The level of expression of lncRNA-HIT in normal breast tissue was low and gradually increased to invasive carcinoma suggesting lncRNA-HIT may play a role in tumor progression in humans (Fig. 9*D*). These data further support the findings demonstrating the involvement of lncRNA-HIT in EMT and invasion as observed in NMuMG and 4T1 cells, and suggest that conserved human lncRNA-HIT could play a pivotal role in breast cancer metastasis.

DISCUSSION

Accumulating studies have demonstrated that the TGFβ pathway plays a critical role in breast cancer metastasis and several protein-coding genes and miRNAs have been described in this process (23–26). Moreover, a recent report demonstrated that lncRNA-ATB is up-regulated in TGFβ-treated SMMC-7721 hepatoma cells and plays a significant role in hepatocellular carcinoma metastasis (27). In this study, we report a lncRNA expression signature of TGFβ-induced EMT in mouse mammary gland epithelial (NMuMG) cells. Over 600 lncRNAs were significantly up-regulated or down-regulated during EMT, respectively. Furthermore, we showed that lncRNA-HIT, one of the top up-regulated lncRNA, plays an important role in TGFβ-induced EMT, cell migration, and invasion. Depletion of lncRNA-HIT can reverse the process of EMT-associated gene expression of E-cadherin and Vimentin. Furthermore, lncRNA-HIT expression is significantly elevated in the highly metastatic cell line 4T1 in comparison to 3 other isogenic mouse cell lines with less metastatic capacity. Depletion of lncRNA-HIT in this cell line results in a significant reduction of migration and invasion as well as lung metastasis and tumor growth. LncRNA-HIT is conserved in sequence and genomic location from mouse to human. Human lncRNA-HIT expression is associated with more invasive tumor and breast cancer progression. These findings are important for several reasons. First, this is the first study to demonstrate an lncRNA expression signature of TGFβ-induced EMT in mouse mammary gland epithelial cells.



Tumor types	lncRNA-HIT Low	lncRNA-HIT High	%	P-value
Non-invasive cancer (n =15)	13	2	13	0.0024
Invasive carcinoma (n = 74)	45	29	39	

FIGURE 9. **Elevated levels of lncRNA-HIT are associated with invasive breast cancer in human.** *A,* lncRNA-HIT is conserved between mouse and human. *B,* locked nucleic acid *in situ* hybridization. LNA-oligonucleotide of lncRNA-HIT was labeled with digoxigenin-ddUTP using the digitonin 5′-end labeling kit and hybridized to a human breast cancer tissue microarray. Representative photomicrographs of sections of a normal, hyperplasia, non-invasive, and invasive breast tumor tissues are shown. *C* and *D*, summary of LNA-ISH results. Overexpression of lncRNA-HIT was found to be more frequently detected in invasive breast carcinoma than non-invasive tumor (*C*) and associated with tumor progression (*D*). *Asterisks* represent $p < 0.05$.

Second, this study established a critical role of a previously uncharacterized lncRNA, lncRNA-HIT, in TGFβ-induced EMT. Finally, conserved lncRNA-HIT was shown to be an important lncRNA in mouse and human breast cancer progression.

LncRNA-HIT resides in the 5′ distal Hoxa gene cluster between but not overlapping with Hoxa13 and Hoxa11-as. This region of the genome is highly conserved across several species and several functional lncRNAs have been demonstrated to be important within the Hoxa clusters, including neighboring HOTTIP and HOXA11-AS (12, 22, 28). HOTTIP has been shown to positively regulate the protein coding genes in close proximity, notable HOXA13, through the recruitment of WDR5, a component of the histone methyltransferase protein MLL complex that promotes transcriptional activation (12). It has also recently been reported that elevated levels of human HOTTIP and HOXA13 expression associates with disease progression and poorer survival in hepatocellular carcinoma (29). However, lncRNAs HOTTIP and HOXA11-AS were not induced by TGFβ in our microarray analysis, suggesting lncRNA-HIT has a unique role in TGFβ-induced EMT. In addition, we demonstrated that E-cadherin promoter activity and the mRNA level were significantly repressed by lncRNA-HIT and that enforced expression of E-cadherin largely abrogated lncRNA-HIT-induced EMT, cell migration, and invasion. These findings suggest that E-cadherin is a key target of lncRNA-HIT, which regulates E-cadherin via *trans* mechanism.

Although microarray analysis revealed expression changes for a number of typical TGFβ-regulated protein-coding genes upon TGFβ-induced EMT, the array did not show altered expression of the lncRNA-HIT neighbor genes including HOXA13 and HOXA11. Further investigation is warranted to determine the mechanism of lncRNA-HIT regulation of E-cadherin, as well as its use as a therapeutic target for breast cancer metastasis intervention.

Acknowledgment—We thank the Molecular Genomics Facility at H. Lee Moffitt Cancer Center for providing microarray and sequencing.

REFERENCES

1. Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., and Chang, H. Y. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129,** 1311–1323
2. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J., and Lee, J. T. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322,** 750–756
3. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., and Young, R. A. (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130,** 77–88
4. Beltran, M., Puig, I., Peña, C., García, J. M., Alvarez, A. B., Peña, R., Bonilla, F., and de Herreros, A. G. (2008) A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev.* **22,** 756–769
5. Schmierer, B., and Hill, C. S. (2007) TGFβ-SMAD signal transduction: molecular specificity and functional flexibility. *Nat. Rev. Mol. Cell Biol.* **8,** 970–982
6. Kong, W., Yang, H., He, L., Zhao, J. J., Coppola, D., Dalton, W. S., and Cheng, J. Q. (2008) MicroRNA-155 is regulated by the transforming growth factor β/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA. *Mol. Cell. Biol.* **28,** 6773–6784

7. Diederichs, S. (2014) The four dimensions of noncoding RNA conservation. *Trends Genet.* **30,** 121–123
8. Johnsson, P., Lipovich, L., Grandér, D., and Morris, K. V. (2014) Evolutionary conservation of long non-coding RNAs: sequence, structure, function. *Biochim. Biophys. Acta* **1840,** 1063–1071
9. Gottardi, C. J., Wong, E., and Gumbiner, B. M. (2001) E-cadherin suppresses cellular transformation by inhibiting β-catenin signaling in an adhesion-independent manner. *J. Cell Biol.* **153,** 1049–1060
10. Yang, H., Kong, W., He, L., Zhao, J. J., O'Donnell, J. D., Wang, J., Wenham, R. M., Coppola, D., Kruk, P. A., Nicosia, S. V., and Cheng, J. Q. (2008) MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN. *Cancer Res.* **68,** 425–433
11. Allred, D. C., Harvey, J. M., Berardo, M., and Clark, G. M. (1998) Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod. Pathol.* **11,** 155–168
12. Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B. R., Protacio, A., Flynn, R. A., Gupta, R. A., Wysocka, J., Lei, M., Dekker, J., Helms, J. A., and Chang, H. Y. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472,** 120–124
13. Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011) The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Mol. Cell* **43,** 1040–1046
14. Yang, F., Huo, X. S., Yuan, S. X., Zhang, L., Zhou, W. P., Wang, F., and Sun, S. H. (2013) Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. *Mol. Cell* **49,** 1083–1096
15. Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., Forrest, A. R., Carninci, P., Biffo, S., Stupka, E., and Gustincich, S. (2012) Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491,** 454–457
16. Hacisuleyman, E., Goff, L. A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., McClanahan, P., Hendrickson, D. G., Sauvageau, M., Kelley, D. R., Morse, M., Engreitz, J., Lander, E. S., Guttman, M., Lodish, H. F., Flavell, R., Raj, A., and Rinn, J. L. (2014) Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* **21,** 198–206
17. Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., Freier, S. M., Bennett, C. F., Sharma, A., Bubulya, P. A., Blencowe, B. J., Prasanth, S. G., and Prasanth, K. V. (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **39,** 925–938
18. Onder, T. T., Gupta, P. B., Mani, S. A., Yang, J., Lander, E. S., and Weinberg, R. A. (2008) Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res.* **68,** 3645–3654
19. Perl, A. K., Wilgenbus, P., Dahl, U., Semb, H., and Christofori, G. (1998) A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature* **392,** 190–193
20. Vleminckx, K., Vakaet, L., Jr., Mareel, M., Fiers, W., and van Roy, F. (1991) Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. *Cell* **66,** 107–119
21. Frixen, U. H., Behrens, J., Sachs, M., Eberle, G., Voss, B., Warda, A., Löchner, D., and Birchmeier, W. (1991) E-Cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *J. Cell Biol.* **113,** 173–185
22. Manzanares, M., Wada, H., Itasaki, N., Trainor, P. A., Krumlauf, R., and Holland, P. W. (2000) Conservation and elaboration of *Hox* gene regulation during evolution of the vertebrate head. *Nature* **408,** 854–857
23. Kang, Y., He, W., Tulley, S., Gupta, G. P., Serganova, I., Chen, C. R., Manova-Todorova, K., Blasberg, R., Gerald, W. L., and Massagué, J. (2005) Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 13909–13914
24. Nam, J. S., Suchar, A. M., Kang, M. J., Stuelten, C. H., Tang, B., Michalowska, A. M., Fisher, L. W., Fedarko, N. S., Jain, A., Pinkas, J., Lonning, S., and Wakefield, L. M. (2006) Bone sialoprotein mediates the tumor cell-targeted prometastatic activity of transforming growth factor β in a mouse model of breast cancer. *Cancer Res.* **66,** 6327–6335
25. Roberts, A. B., and Wakefield, L. M. (2003) The two faces of transforming growth factor beta in carcinogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 8621–8623
26. Xie, W., Mertens, J. C., Reiss, D. J., Rimm, D. L., Camp, R. L., Haffty, B. G., and Reiss, M. (2002) Alterations of Smad signaling in human breast carcinoma are associated with poor outcome: a tissue microarray study. *Cancer Res.* **62,** 497–505
27. Yuan, J. H., Yang, F., Wang, F., Ma, J. Z., Guo, Y. J., Tao, Q. F., Liu, F., Pan, W., Wang, T. T., Zhou, C. C., Wang, S. B., Wang, Y. Z., Yang, Y., Yang, N., Zhou, W. P., Yang, G. S., and Sun, S. H. (2014) A long noncoding RNA activated by TGF-β promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* **25,** 666–681
28. Chau, Y. M., Pando, S., and Taylor, H. S. (2002) HOXA11 silencing and endogenous HOXA11 antisense ribonucleic acid in the uterine endometrium. *J. Clin. Endocrinol. Metab.* **87,** 2674–2680
29. Quagliata, L., Matter, M. S., Piscuoglio, S., Arabi, L., Ruiz, C., Procino, A., Kovac, M., Moretti, F., Makowska, Z., Boldanova, T., Andersen, J. B., Hämmerle, M., Tornillo, L., Heim, M. H., Diederichs, S., Cillo, C., and Terracciano, L. M. (2014) Long noncoding RNA HOTTIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. *Hepatology* **59,** 911–923