The role of the N6-methyladenosine RNA modification in mediating HOTAIR-protein interactions

A Thesis

Presented To

The Faculty of the Department of Molecular Biology

Colorado College

In Partial Fulfillment

of the Requirements for the Degree of

Bachelor of Arts

by

Madeline Chrupcala

May 2019

ABSTRACT

The long noncoding RNA (lncRNA) HOTAIR acts in *trans* to epigenetically silence a 40-kb region of the *HOXD* gene cluster during development. In aggressive breast cancer, HOTAIR overexpression promotes metastasis, drug resistance, recurrence, and is a negative prognostic factor. However, the precise molecular mechanisms by which HOTAIR induces the formation of heterochromatin at specific sites in the genome to silence tumor suppressor genes remains widely unknown. To learn more about the molecular role of HOTAIR, we explored whether it was modified by the N6-methyladenosine RNA modification, which is dysregulated in breast cancer and has been proposed to play an essential role in gene repression by mediating lncRNA-protein interactions. Because we mapped m6A to a single nucleotide within the second domain of HOTAIR, we performed a quantitative proteomic analysis to identify additional HOTAIR-protein interactions that may be mediated by the post-transcriptional modification. We found that HOTAIR interacts with the YTH domain-containing protein 1 (YTHDC1), a member of a family of proteins known to recognize m6A residues. To determine the role of m6A in the HOTAIR-YTHDC1 interaction, we performed RNA pulldown assays with purified YTHDC1 and *in vitro* transcribed wild-type and m6A mutant HOTAIR in the presence or absence of methylation. Our data found that YTHDC1 preferentially interacts with wild-type methylated HOTAIR, suggesting that m6A mediates the HOTAIR-YTHDC1 interaction. However, it is not yet known if the m6A residue of HOTAIR recruits YTHDC1 to promote the transcriptional silencing of important suppressor genes. Thus, future research exploring the role of the m6A-mediated HOTAIR-YTHDC1 interaction in metastatic breast cancer will demonstrate whether a better understanding of the lncRNA can drive the development of breast cancer therapeutics. This is of utmost importance because invasive breast cancer remains a leading cause of death for women worldwide despite recent advancements in diagnosis and treatments.

INTRODUCTION

Long noncoding RNAs are key regulators of gene expression.

Over the past decade, advancements in high-throughput sequencing have shown that while more than 80% of the human genome contains functional DNA elements, only approximately 2% of the genome demonstrates protein-coding potential [1-9]. The Encyclopedia of DNA Elements (ENCODE) Consortium allowed for the identification of previously undescribed regions of the genome including long noncoding RNAs (lncRNAs). LncRNAs stem from tens of thousands of RNA polymerase II transcription events, are longer than 200 nucleotides, and undergo capping, polyadenylation, and splicing [7-9]. Thus, except for their inability to code for proteins, lncRNAs fully resemble messenger RNAs (mRNAs). Some lncRNAs are believed to have conserved function and secondary structure, despite having little overall sequence conservation across species [7]. Although their exact molecular and physiological function is not entirely understood, it is becoming increasingly evident that lncRNAs are major regulators of chromatin structure, gene expression, differentiation, and development [1-11].

lncRNAs interact with chromatin-modifying complexes to regulate chromatin structure.

Chromatin is a dynamic macromolecular structure, varying in its degrees of compaction to promote the formation of either transcriptionally active euchromatin or transcriptionally inactive heterochromatin [2,7,9,11-17]. Notably, ‘facultative heterochromatin’ can switch between the active and inactive chromatin states through a tightly regulated mechanism. The formation of facultative heterochromatin represents a selective, cell-specific event that enables the expression of specific genes depending on the developmental time frame [2]. Recent studies have shown lncRNAs influence both facultative and constitutive heterochromatin through direct interactions with chromatin-modifying complexes, including methyltransferases, demethylases, acetyltransferases, and deacetylases. These interactions induce chemical modifications to

histone proteins or DNA nucleotides that regulate transcription by exposing different regions of DNA to transcription factors and RNA polymerase II [7].

To examine the interactions between lncRNAs and chromatin-modifying complexes, a recent study performed a hybridized RNA immunoprecipitation-chromatin immunoprecipitation (RIP-ChIP) assay using antibodies against the Polycomb repressive complex 2 (PRC2) [18]. They focused on PRC2, a methyltransferase that tri-methylates H3K27 to repress transcription, because it was previously proposed to regulate the epigenetic landscape at specific sites in the genome [18]. This method enabled them to identify physical interactions between PRC2 and numerous lncRNAs. Because PRC2 promotes H3K27 tri-methylation, these findings support a model in which lncRNAs regulate chromatin structure by recruiting protein complexes to induce chemical modifications to histone proteins [18]. Moreover, siRNA depletion of the lncRNAs they identified to physically interact with PRC2 resulted in changes in gene expression profiles. It is therefore likely that these lncRNAs recruit PRC2 to induce epigenetic modifications to ultimately regulate the transcription of specific genes. It is important to note that the methylation of lysine residues on histone tails does not alter the charge of the histone protein even though multiple methylation states exist [19]. Therefore, it is not surprising that the mono-, di-, or tri-methylation of lysine residues is responsible for significant added complexity to the regulation of gene expression that is mediated by a number of methyltransferases and demethylases.

Recent studies have shown that the DNA methylation machinery sometimes requires lncRNAs [20]. For instance, the lncRNA *TCF21* antisense RNA inducing demethylation (TARID) has been shown to bind the growth arrest and DNA damage-inducible protein (GADD45α), resulting in the recruitment of TDG and TETs proteins and the demethylation of the *TCF21* promoter [21-22]. Because methylation silences gene expression, the lncRNA TARID enables the transcription of the tumor suppressor gene transcription factor 21 (*TCF21*) by reversing the modification.

Furthermore, the specificity of lncRNA expression to certain cell types, tissue types, and developmental time frames might explain how demethylating enzymes are targeted to specific sites in the genome. Consistent with this hypothesis, the demethylating part of TARID was shown to interact with the *TCF21* promoter, suggesting a gene-specific demethylation mechanism [21]. Taken together, these data demonstrate the ability of lncRNAs to recruit chromatin-modifying complexes to affect gene transcription.

lncRNAs are dysregulated in cancer.

It is now well-understood that cancer is a disease of the cell cycle, driven by mutations that disrupt gene expression patterns and protein function [24-25]. However, it is becoming increasingly clear that lncRNAs that induce epigenetic modifications to alter chromatin dynamics also play integral roles in the initiation and progression of cancer [2-8, 12-13, 19-25]. Indeed, the dysregulation of lncRNAs has been shown to repress tumor suppressor genes and/or activate oncogenes to promote tumorigenesis. For instance, the antisense noncoding transcript p15AS has been shown to repress p15, a cyclin-dependent kinase inhibitor and well-known tumor suppressor [23]. p15AS expression increased H3K9 demethylation and decreased H3K4 demethylation in the exogenous *p15* promoter region, the exogenous *p15* exon 1 region, and the endogenous sequence around the *p15* transcriptional start site, suggesting that the antisense RNA triggers the formation of heterochromatin to promote the transcriptional repression of *p15*. Like many other lncRNAs, p15AS is overexpressed in numerous cancers, including acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL). The differential expression of lncRNAs in cancer suggests that they may serve as potential powerful therapeutic targets. Indeed, the expression of lncRNAs can be suppressed by RNA interference (RNAi) technology, antisense oligonucleotides, and small molecule inhibitors, suggesting that targeting cancer-associated lncRNAs might be an effective treatment option [24]. However, there are many challenges that need to be addressed before therapeutic options that target

cancer-associated lncRNAs can be established. For instance, it will be crucial to understand the detailed molecular mechanisms that lncRNAs employ to regulate gene expression. It will also be extremely important to determine the subsequent off target effects of targeting a specific lncRNA.

HOTAIR functions *in trans* to silence target genes.

One well-studied lncRNA that is implicated in cancer is the HOX Transcript Antisense Intergenic RNA (HOTAIR), a ~ 2.2-kb transcript derived from the *HOXC* gene cluster between *HOXC11* and *HOXC12* on human chromosome 12q13.13 [15]. HOTAIR was the first lncRNA identified to act *in trans* to epigenetically silence a 40-kb region of the *HOXD* gene cluster, including the *HOXD8*, *HOXD9*, *HOXD10*, and *HOXD11* genes that encode proteins that control embryonic development [11-15, 26, 29]. HOTAIR, like many other lncRNAs, has been postulated to modulate chromatin structure through reversible modifications to histone proteins that regulate the expression of target genes. Indeed, HOTAIR knockdown resulted in a significant loss of H3K27-trimethylation at the HOXD locus as detected by ChIP-chip analyses [2]. H3K27me3, a histone modification that signals the formation of constitutive or facultative heterochromatin to silence gene expression, is induced by the histone methyl-transferase EZH2 (Enhancer of Zeste homolog 2). EZH2 along with SUZ12 (suppressor of zeste 12 homolog), EED (embryonic ectoderm development), and RbAp46/48 (retinoblastoma binding protein 46/48) form the methyltransferase complex PRC2 [2]. Along with the loss of H3K27me3, HOTAIR knockdown also resulted in the loss of occupancy of SUZ12 at the HOXD locus. Coupled with the large-scale RIP-ChIP hybridized assay mentioned previously, these findings suggest that HOTAIR physically interacts with PRC2 and guides it to specific sites in the genome to silence gene expression through reversible modifications to histone proteins [2, 18].

The involvement HOTAIR in epigenetic regulation of gene expression was further solidified by the finding that the lncRNA interacts with the histone demethylase LSD1 (lysine specific demethylase 1A) [2]. LSD1, a flavin-dependent monoamine oxidase, forms a protein complex with REST (RE1-Silencing Transcription factor) and CoREST, both of which are extremely important in silencing gene expression. These studies demonstrate that HOTAIR contains distinct binding domains for PRC2 and LSD1 [26]. The 5' end of HOTAIR binds PRC2 to induce histone H3K27me3 at specific target sites, whereas the 3' end binds LSD1 to demethylate histone H3 at lysine 4 [2,9]. It is therefore likely that HOTAIR functions as a molecular scaffold for both histone-modifying complexes. In this model, HOTAIR regulates gene expression influences through the recruitment of both LSD1 and PRC2 to induce histone modifications that signal the formation of heterochromatin at target loci.

PRC2 is dispensable for HOTAIR-mediated gene repression.

In contrast to the studies above, both PRC2 and its cofactor JARID2 have been shown to bind multiple RNAs with a relatively high affinity [28]. Indeed, the inhibition of transcription was sufficient to trigger the recruitment of PRC2 to many sites in the genome [27]. It is therefore likely that PRC2 binding is dependent on size of the RNA rather than on the intricate secondary structures [27]. Although PRC2 readily binds many RNAs, only a subset of these binding events resulted in the induction of H3K27 tri-methylation at distinct sites in the genome [26-27]. These findings demonstrate that the methyltransferase activity of PRC2 requires a unique context, suggesting that the lncRNA itself likely plays an important role in silencing the expression of target genes. Moreover, HOTAIR was shown to maintain its repressive function in cells that lacked essential components of PRC2 [29]. In light of these findings, it is reasonable to speculate that some of the observed changes in chromatin dynamics following the recruitment of HOTAIR are a secondary consequence of HOTAIR-mediated gene repression. Thus, it will be

extremely important to identify additional factors that interact with HOTAIR to silence the expression of target genes.

HOTAIR is overexpressed in breast cancers.

The upregulation of HOTAIR has been implicated in a variety of cancers, including metastatic breast cancer, small cell lung carcinoma, glioma, hepatocellular carcinoma, gastric cancer, and pancreatic cancer [2-9, 12-13, 23, 25, 30-32]. Because cancer is frequently characterized by the aberrant expression of genes, a study aimed to determine whether HOTAIR overexpression influences the methyltransferase activity of PRC2. To do this, they performed ChIP experiments with antibodies against EZH2, H3K27me3, and SUZ12 followed by hybridization to a genome-wide promoter array [13]. The results from these experiments demonstrate that HOTAIR overexpression can lead to the genome-wide retargeting of PRC2 to 854 new genes, including the progesterone receptor (*PRG1*), cell adhesion molecules of the protocadherin (*PCDH*) gene family, and the junctional adhesion molecule-2 (*JAM2)*. Because these genes have demonstrated an ability to restrict tumor growth and metastasis, these results suggest that HOTAIR contributes to breast cancer pathogenesis by altering gene expression profiles to promote tumor growth, metastasis, invasion and migration. In addition, siSENSE-mediated HOTAIR knockdown upregulated the expression of tumor suppressor and metastasis restricting genes, including *HOXD10* and *PCDH10*, and induced apoptotic cell death in MCF-7 breast cancer cells [2]. Given the general lack of understanding of how HOTAIR contributes to tumorigenesis, additional research exploring the mechanisms behind HOTAIR-mediated repression of HOTAIR tumor suppressor genes will demonstrate whether a better understanding of the lncRNA can drive the development of novel cancer therapeutics.

The N6-methyladenosine RNA modification machinery is dysregulated in cancer.

Eukaryotes rely on changes in gene expression for proper development and homeostasis [35-42]. These changes in gene expression are dependent on the synchronization of multiple levels of regulation, including transcriptional, post-transcriptional, and post-translational control [35-42]. At the post-transcriptional level, chemical modifications to RNA regulate gene expression by affecting multiple aspects of RNA transcripts and processing events [36-43]. For instance, N6-methyladenosine (m6A), the most common post-transcriptional modification in mRNAs and lncRNAs, is proposed to regulate RNA stability, splicing, transport, translation, secondary structure, and RNA-protein interactions through unknown mechanisms [35-43]. The relative abundance of m6A on RNA is determined by the dynamic interactions between its methyltransferases and demethylases. The reversible modification is established by a multi-protein methyltransferase complex consisting of Methyltransferase Like 3 (METTL3), METTL14, Wilms Tumor 1 Associated Protein (WTAP), KIAA1429, RNA Binding Motif Protein 15 (RBM15), and zinc finger CCHH domain-containing protein 13 (ZC3H13) [24, 40]. The modification is effectively removed by the fat mass and obesity-associated protein (FTO) and AlkB homolog 5 RNA demethylase (ALKBH5) that demethylate the adenosine residue.

Given the regulatory role of m6A in RNA processing events, it is unsurprising that the misregulation of the post-transcriptional modification has been implicated in various human diseases, including cancer [25, 35-43]. It is believed that the aberrant expression of the m6A modification machinery promotes cancer metastasis, growth, and survival through diverse mechanisms. For instance, a previous study found that FTO is significantly downregulated in estrogen-receptor-positive (ER+), estrogen-receptor-negative (ER-), progesterone-receptor-positive (PR+), progesterone-receptor-negative (PR-), human epidermal growth factor receptor 2- positive (HER2+), and human epidermal growth factor receptor 2-negative (HER-) breast cancers [44]. Because similar results were obtained for all subtypes, they concluded that breast tumor tissues have significantly less demethylase activity regardless of their hormone receptor

status. The dysregulation of m6A machinery in cancer suggests that targeting the post-transcriptional modification might represent a powerful anti-cancer target. Indeed, a recent study has shown that inhibiting the active subunit of the m6A methyltransferase complex decreased proliferation, migration, and colony formation in hepatocellular carcinoma (HCC) [45]. Additionally, the *in vivo* knockdown of METTL3 repressed HCC tumorigenesis and lung metastasis [40, 44]. Taken together, these data establish a connection between the disruption of m6A methylation-dependent pathways, changes in gene expression profiles, and the development of aggressive cancers.

m6A promotes XIST-mediated gene repression.

XIST (X-inactive specific transcript), a highly methylated lncRNA, mediates the transcriptional silencing of one X-chromosome during early embryonic development to balance X-linked gene expression between male and female mammals [47]. Notably, XIST was unable to silence gene expression in cells treated with siRNA to knockdown METTL3 activity indicating that the m6A modification is required for XIST-mediated X-chromosome inactivation (XCI) [47]. To explore the mechanism by which m6A promotes XIST function, this same study performed knockdown experiments of YTH proteins, a family of proteins known to recognize m6A residues. YTHDC1 (alternative name YT521-B) is the only YTH protein localized to the nucleus and has been proposed to regulate RNA processing events. In light of these findings, it is unsurprising that the knockdown of only YTHDC1 inhibited XIST-mediated gene repression. Furthermore, the artificial tethering of YTHDC1 to XIST induced the repressive function of the lncRNA in the absence of the methyltransferase complex. Together, these discoveries propose a model in which the m6A residues of XIST recruit YTHDC1 to regulate X-chromosome expression during early female embryonic development. The ability of m6A to recruit proteins to confer downstream effects on target genes further signifies the importance of chemical modifications to lncRNAs in gene expression regulation. Because the post-transcriptional modification has

previously been shown to influence RNA secondary structure, it is plausible that both direct and indirect interactions between m6A modified lncRNAs and their subsequent reader proteins occur. Thus, deciphering the contribution of m6A to specific lncRNA transcripts may lead to a better understanding of the roles of different modified lncRNAs and their reader proteins in regulating gene expression.

HOTAIR contains the m6A modification.

The discovery that m6A promotes XIST-mediated X-chromosome inactivation raised the question of whether the modification influences the function of additional lncRNAs. Although the dysregulation of the m6A pathway and HOTAIR have previously been shown to promote increased metastasis in breast cancer, the two mechanisms are not known to interact. Thus, we were interested in exploring the possibility of the post-transcriptional modification in HOTAIR and whether it may mediate interactions with unidentified proteins. Using a novel method termed m6A eCLIP (meCLIP), we were able to detect the m6A modification to a single nucleotide within the second domain of HOTAIR in epithelial-like MCF-7 breast cancer cells (ER+, PR+, HER2-), which express endogenously elevated levels of the lncRNA. The results from these studies reveal that HOTAIR and many of its target mRNAs are m6A modified.

In the present study, we set out to further investigate the link between HOTAIR and m6A in metastatic breast cancer. Given the ability of the m6A RNA modification to recruit specific reader proteins to confer its downstream effects on gene expression, we were interested in identifying additional HOTAIR-protein interactions. We performed RNA pulldown experiments to screen for interactions with proteins known to recognize m6A residues. From this, we identified the interaction with YTHDC1, which contributes to XIST-mediated transcriptional repression of an X-chromosome. Furthermore, additional RNA pulldown experiments with WT m6A methylated, WT m6A unmethylated, m6A mutant methylated, and m6A mutant unmethylated

HOTAIR revealed that YTHDC1 preferentially binds to WT m6A methylated HOTAIR. Our studies present a potential new m6A-dependent mechanism for HOTAIR-mediated gene repression.

MATERIALS AND METHODS

Competent cell transformation

Competent DH5α *Escherichia coli* cells were incubated with 50ng of pcDNA-FLAG-YTHDC1 for 30 min, on ice. Afterwards, cells were heat shocked for 2.5 min at 37°C, immediately transferred to an ice bath for 1.5 min, shaken at 37°C with LB broth, and spread on agar plates with 100µg/mL Carbenicillin and incubated at 37°C overnight. A single transformed bacteria colony was shaken at 37°C overnight in LB-Ampicillin. Approximately 400µg of pcDNA-FLAG-YTHDC1 plasmid DNA was isolated from bacteria culture using the ZymoPure II Plasmid kit (Zymo Research). Plasmid DNA concentrations were measured with a Nanodrop Spectrophotometer (Thermo Scientific).

Cell culture and transfections

Human embryonic kidney 293T (HEK293T) cells were maintained in Dulbecco's Modified Eagle's Media (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% Penicillin-Streptomycin (Life Technologies) at 37°C in a humidified incubator with 5% CO_2. 48µg of isolated pCDNA-FLAG-YTHDC1 plasmid DNA and Lipofectamine 2000 (Invitrogen) were diluted in Opti-MEM media (Gibco) and incubated together (1:1) for 20 min at room temperature to form a DNA-lipid complex. HEK293T cells at 75-80% confluency were transfected with the DNA-lipid complex and incubated at 37°C for 48 hours-72 hours. The transfected cell lysate was harvested, pelleted, and frozen in 1X Phosphate Buffered Saline (PBS, pH 7.4) at -80°C until further use.

In vitro transcription of RAT-tagged HOTAIR

PCR with the Phusion polymerase was used to add 5' T7 promoter sequences and a RAT tag to generate templates for *in vitro* transcription of the m6A-containing domain of HOTAIR domain 2, MW= $1.7x10^5$ KDa, using pLPCX HOTAIR and pBabePuro-HOTAIR dm6A plasmids with resistance to ampicillin. PCR was performed with the appropriate primers (Table 1) and purified with the E.Z.N.A. Cycle Pure Kit (Omega Biotek). *In vitro* transcription of WT and m6A-mutant (A→ T) HOTAIR was performed using the MEGAscript T7 transcription kit (Invitrogen) for 4 h at 37°C. RNA was treated with DNase1 (Ambion) for 15 min and purified with the RNeasy kit (QIAGEN). RNA concentrations were measured with a Nanodrop spectrophotometer (Thermo Scientific). *In vitro* transcripts were analyzed by gel electrophoresis.

Anti-FLAG M2 affinity purification

The harvested HEK293T cell pellet was resuspended in 1X lysis buffer (1M Tris-HCl pH 7.4, 5M NaCl, 500mM EDTA, 20% Triton-X, and Protease Inhibitors) for 20 min at 4°C. Anti-FLAG M2 affinity resin were equilibrated with 1X lysis buffer at 500xg for 2 min at 4°C and incubated with the cell lysate for 4 hours at 4°C with nutation. After the incubation period, the resin was transferred to low retention tubes and washed with 1X lysis buffer. FLAG-tagged YTHDC1 protein was eluted off the anti-FLAG M2 affinity resin by competitive elution with either 1X wash buffer or 1X lysis buffer and 0.1mg/ml 3X FLAG peptide at 1000rpm for 10 min at room temperature. The supernatant was saved for 3 separate elutions and analyzed using SDS-PAGE and western blotting.

Methyltransferase activity assays

4.25µg of WT and m6A mutant HOTAIR RNA were methylated with 500µM S-Adenosyl methionine (SAM), 5µM METTL3/14, and 1X Methyltransferase buffer (diluted from 10X stock: 100mM DTT, 1M Tris-HCl, pH 7.5, and 20% Triton-X). 4.25µg of WT and m6A mutant HOTAIR

RNA were incubated with 500μM SAM, 1X, and 10X methyltransferase buffer to generate unmethylated RNA samples. The unmethylated and methylated RNA was purified using the RNeasy kit (Qiagen) and the RNA concentrations were measured with a Nanodrop Spectrophotometer (Thermo Scientific).

RAT-tagged HOTAIR-YTHDC1 pulldown

0.8 μg of wild type and m6A mutant HOTAIR RNA were incubated with LexA-PP7 fusion protein and Hypotonic Lysis buffer (HLB; 20mM Hepes, pH 7.9, 300mM NaCl, 2mM MgCl2, 0.1% NP-40, 10%, Glycerol, 0.1mM PMSF, 0.5 DTT, RNase Inhibitor) for 30 min at 25°C to prebind. Rabbit IgG coupled magnetic dynabeads were washed in HLB-100 buffer and conjugated to the RNA-PP7 complex for 1 hour at 25°C at 1350rpms. The conjugated beads were washed with binding buffer (10mM Hepes, pH 7.4, 150mM KCL, 3mM MgCl, 2mM DTT, 0.5% NP-40, 10% glycerol, 1mM PMSF, and Protease Inhibitors) and nutated with purified YTHDC1 protein for 2.5 hours at 4°C. The beads were washed with wash buffer (20mM Tris pH 7.4, 200mM NaCl, 2mM MgCl, 1mM DTT) and resuspended in 1X SDS loading buffer. Samples were boiled for 5 min at 95°C and analyzed using SDS-PAGE and western blotting.

Western blot analysis

Proteins were separating using SDS-polyacrylamide gel electrophoresis (12% gels) and transferred to nitrocellulose membranes (Millipore). The filters were probed with the indicated primary and secondary antibodies conjugated to horseradish peroxidase. The following antibodies were diluted in 2.5-5% milk with 1X TBST according to the manufacturer's instructions: FLAG M2 (1:1500; Sigma F1804), Anti-mouse HRP (BioRad 170-6516), YTHDC1 (1:500; Abcam 122340), and Anti-rabbit HRP (Biorad 170-5046). Immobilon Chemiluminescent HRP Substrate from Millipore was used for detection using a BioRad ChemiDoc system. Statistical analysis was performed using the Bio-Rad analysis software.

RESULTS

HOTAIR and many of its target genes contain the RNA modification m6A

Given the lack of understanding of the observed gene silencing effect of HOTAIR, we wanted to explore its mechanism of action and, whether it involves the post-transcriptional modification m6A. Because the dysregulation of both HOTAIR and m6A have been shown to promote increased metastatic potential of breast cancer, we hypothesized that the modification may promote the repression of tumor suppression genes. Consistent with this hypothesis, The Cancer Genome Atlas (TCGA) demonstrated that the enzymes that establish and remove the m6A modification are dysregulated in breast cancer. For instance, FTO, the m6A demethylase, is downregulated in breast cancer [44]. These data strongly suggest that breast cancer tissue have higher levels of m6A methylation when compared to normal breast tissue. Furthermore, the m6A writer protein, METTL14, appears to be upregulated in breast cancer [45]. To determine if HOTAIR is m6A modified, we performed m6A eCLIP (meCLIP) in MCF-7 breast cancer cells (ER+, PR+, HER2-), which express endogenously elevated levels of HOTAIR. meCLIP, a novel method development from modified miCLIP and eCLIP protocols, mapped the m6A modification to a single nucleotide in epithelial-like breast cancer cells (Figure 1A). The m6A site in HOTAIR was mapped to the second domain of HOTAIR to nucleotide A783 in the transcript. The presence of m6A in HOTAIR in MCF-7 breast cancer cells was confirmed with m6A-immunoprecipitation (meRIP) using an anti-body against m6A followed by qRT-PCR (Figure 1B). These experiments were also performed in MDA-MB-231 triple-negative breast cancer cells overexpressing HOTAIR, which was also shown to contain the m6A modification (Figure 1C). Notably, MDA-MB-231 cells express significantly lower levels of HOTAIR compared to MCF-7 cells. However, MDA-MB-231 triple-negative breast cancer cells are more metastatic.

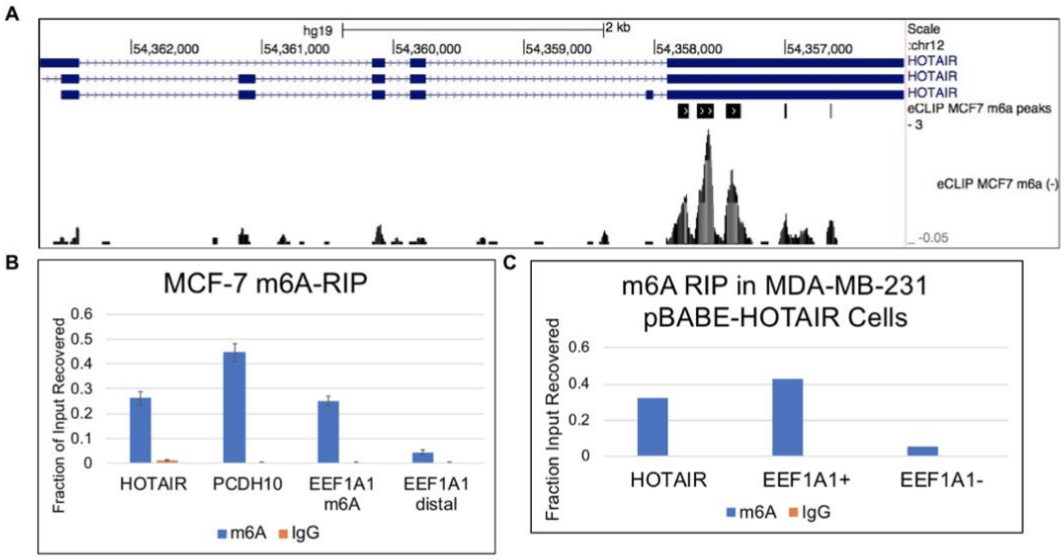

Figure 1. HOTAIR is m6A modified (A) Results from m6A eCLIP (meCLIP) experiments performed in MCF-7 breast cancer cells (estrogen-receptor-positive, progesterone-receptor-positive, HER2-negative). (B) RNA immunoprecipitation (RIP) with antibody against m6A (blue bars) or no antibody (IgG; orange bars) from MCF-7 breast cancer cells. HOTAIR, PCDH10, and EEF1A1 levels were measured with RT-qPCR. (C) RNA immunoprecipitation with antibody against m6A (blue bars) or no antibody (IgG; orange bars) from MDA-MB-231 breast cancer cells (estrogen-receptor-negative, progesterone-receptor-negative, HER2-negative). HOTAIR, EEF1A1+, and EEF1A1- levels were measured with RT-qPCR. These data were generated from experiments performed by Allison Porman and Justin Roberts.

HOTAIR interacts with YTHDC1.

Because the mechanism by which HOTAIR recognizes its targets and triggers the formation of heterochromatin remain elusive, we sought to identify additional interaction partners for the lncRNA that may promote gene repression. Our previous meCLIP and m6A RNA immunoprecipitation experiments indicated that nucleotide A783 in the second domain of HOTAIR is m6A methylated. We reasoned that if the m6A modification on HOTAIR mediates lncRNA-protein interactions to promote HOTAIR-mediated gene repression, then a large-scale proteomic analysis of the entire lncRNA interactome would likely identify interactions with known m6A reader proteins. We focused our efforts on the interaction with YTHDC1, a member of a family of proteins known to recognize the m6A residues on XIST to promote the random inactivation of one X-chromosome during early female development [46]. We performed RNA immunoprecipitation experiments using an antibody against YTHDC1 followed by qPCR. Because the observed HOTAIR cancer phenotypes are primarily due to its overexpression, we

performed these experiments in MCF-7 breast cancer cells (ER+, PR+, HER2-) breast cancer cells, which express the greatest levels of endogenous HOTAIR. As expected, we detected an interaction between HOTAIR and YTHDC1 in epithelial-like MCF-7 breast cancer cells (Figure 2A). Furthermore, we identified an interaction between YTHDC1 and m6A modified PCDH10, an mRNA that has previously been shown to be targeted by HOTAIR (Figure 2A).

The upregulation of HOTAIR has previously been shown to promote increased metastasis in triple-negative breast cancer cells [2-3, 7-9, 12-14]. In light of these findings, we performed the same RNA immunoprecipitation experiments in MDA-MB-231 triple-negative breast cancer cells. Although MDA-MB-231 cells express endogenously lower levels of HOTAIR than MCF-7 cells, they demonstrate less epithelial-like properties and are therefore considered more metastatic. Our data revealed an interaction between HOTAIR and YTHDC1 in metastatic triple-negative breast cancer (Figure 2B). Collectively, the results from the RNA immunoprecipitation experiments in both breast cancer cells indicate that HOTAIR and many of its target mRNAs are m6A modified and bound by YTHDC1. However, the role of the m6A modification in mediating the HOTAIR-YTHDC1 interaction remains elusive.

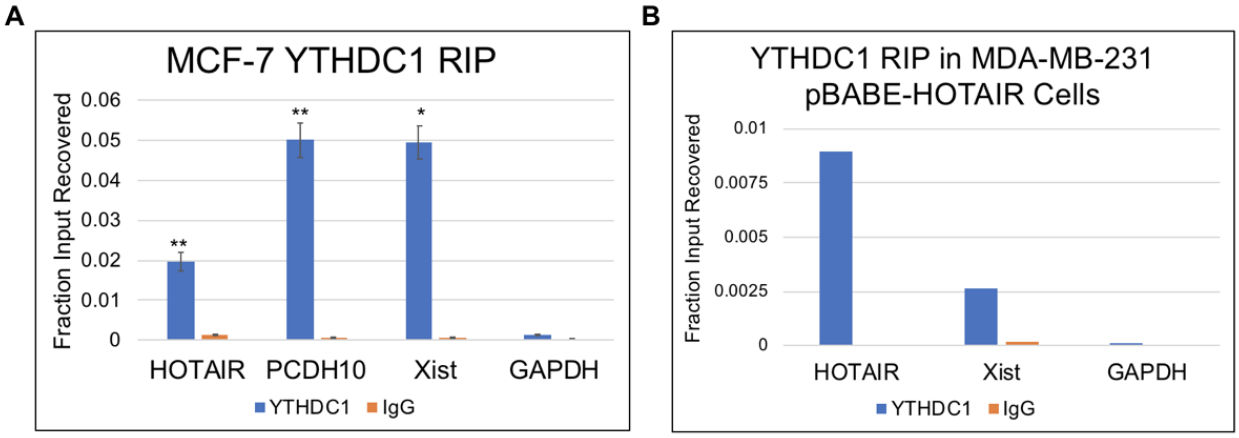

Figure 2. HOTAIR interacts with the m6A "reader" protein YTHDC1. (A) RNA immunoprecipitation with an antibody against YTHDC1 (blue bars) or no antibody (IgG; orange bars) in MCF-7 (ER+, PR+, HER2-). breast cancer cells. HOTAIR, PCDH10, Xist, and GAPDH RNA levels were measured with qPCR. (B) RNA immunoprecipitation with an antibody against YTHDC1 (blue bars) or no antibody (IgG; orange bars) in MDA-MB-231 (triple-negative) breast cancer cells. HOTAIR, Xist, and GAPDH RNA levels were measured with qPCR. These experiments were performed by Allison Porman at the University of Colorado Anschutz Medical Campus.

Failure to elute YTHDC1 off the anti-FLAG M2 affinity resin.

The results from our previous experiments revealed an interaction between HOTAIR and YTHDC1 in breast cancer cells that overexpress the lncRNA. To further probe the HOTAIR-YTHDC1 interaction, we transfected confluent HEK293T cells with FLAG-tagged YTHDC1 cDNA using Lipofectamine 2000. After 48-72 hours, the transfected cells were harvested and FLAG-tagged YTHDC1 protein was purified from the cellular lysates using anti-FLAG M2 affinity resin that contained a mouse monoclonal antibody covalently attached to the resin. Because the antibody selectively bound the FLAG-tagged YTHDC1, we were able to maintain the resin-YTHDC1 interaction while washing away the unspecific proteins in the cellular lysate using a lysis buffer. We ran a sample (bead slurry 1) of the resin before the elutions and were unable to detect the presence of YTHDC1. These results suggest that the FLAG tag of YTHDC1 was unable to bind the monoclonal antibody attached to the resin. However, the presence of YTHDC1 on bead slurry 2, which was a sample of the resin following the purification scheme, indicates that we were unable to elute the protein with a 3X FLAG peptide and wash buffer (Figure 3A). This discrepancy is likely explained by our attempt to pipette a small volume of bead slurry 1 to minimize the effects on the rest of the purification scheme. It is plausible that the bead slurry 1 sample analyzed by gel electrophoresis consisted of buffer rather the resin. With regards to eluting the protein, it is likely that the peptide was unable to outcompete the protein for binding the resin because YTHDC1 was bound with too high of an affinity. Because protein purifications can be extremely sensitive to experimental conditions, we wanted to test if using the lysis buffer throughout the entire purification scheme would be more effective in eluting the protein. We repeated the affinity purification using the 1X lysis buffer and 3X FLAG peptide to elute YTHDC1 and detected stronger bands at around 20-25 kDa and lighter bands at around 50-60 kDa (Figure 3B). These nonspecific bands likely represent the heavy and light chains of the M2 antibody that was not conjugated to the resin. Furthermore, we detected a trace amount of YTHDC1 in bead slurry 1 indicating that the FLAG tag successfully bound the

M2 antibody on the resin. However, we were unable elute the protein off the anti-FLAG M2 affinity resin. We considered introducing harsher elution conditions, including the addition of Urea into the elution buffer. However, we wanted to maintain a functional protein so that we could continue with our interaction analysis. In the interest of time and our experimental aims, we continued using the unbound lysate that was shown to contain the YTHDC1 protein. While the unbound lysate contains unspecific proteins that could interfere with the HOTAIR-YTHDC1 interaction, we reasoned that the lysate contained a sufficient amount of YTHDC1 for successful RNA pulldown experiments.

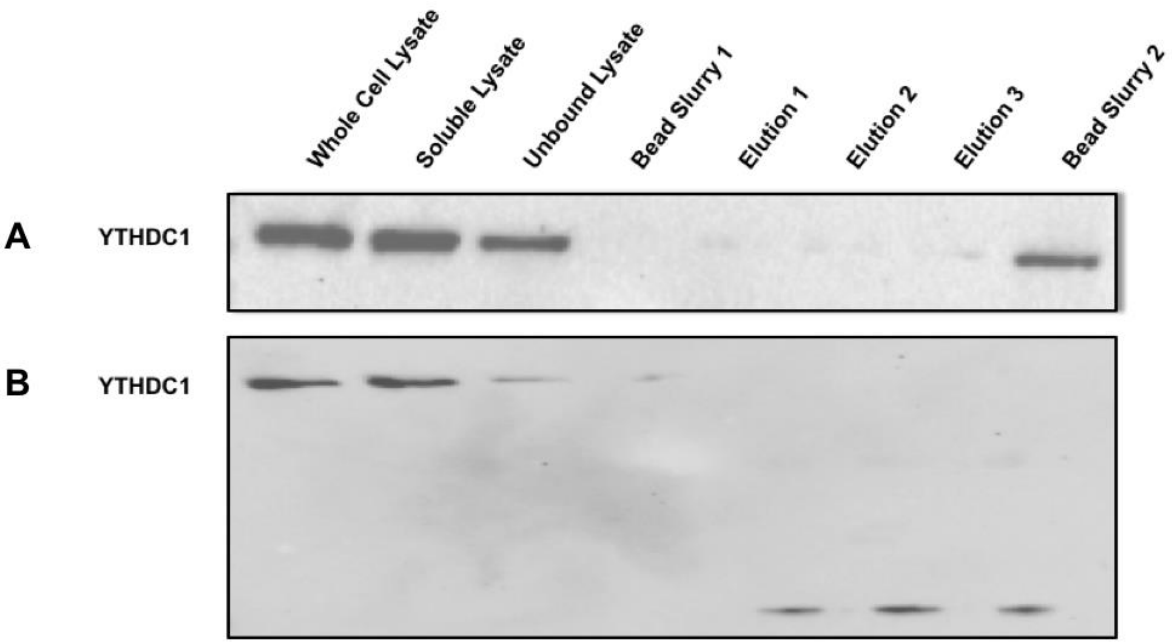

Figure 3. 3X FLAG peptide failed to elute FLAG-tagged YTHDC1. 293T cells were transfected with 48µg of pCDNA-FLAG-YTHDC1 using Lipofectamine 2000 and opti-MEM medium. Cells were harvested after 72 hours and purified using anti-FLAG M2 affinity resin and (A) 1X lysis buffer or (B) 1X wash buffer. Affinity purifications were analyzed with western blotting using an antibody against the FLAG tag.

m6A modified and m6A mutant HOTAIR were *in vitro* transcribed.

Because we wanted to test whether the m6A modification mediates the HOTAIR-YTHDC1 interaction, we generated *in vitro* transcripts of the second domain of wild-type and m6A mutant HOTAIR. Although the secondary structure of HOTAIR is complex and consists of four

molecular domains, we focused our efforts on the second domain that we had previously shown to contain an m6A residue. To generate the m6A mutant, we introduced a point mutation in which the Adenine was substituted to a Thymine in the m6A mutant HOTAIR. Two step PCRs were used to generate PCR-amplified DNA (Figure 4; Left and Middle), which were then used to generate RAT-tagged *in vitro* transcripts of WT and m6A mutant HOTAIR (Figure 4; Right).

To further characterize the role of m6A in facilitating the HOTAIR-YTHDC1 interaction, we incubated the transcripts with S-Adenosyl methionine, the enzymatic cofactor that contributes to most methylation reactions within the cell, and the m6A methyltransferase complex: METTL3 and METTL14. This incubation was sufficient to establish the m6A modification. Because we wanted to determine whether YTHDC1 preferentially interacts with m6A modified HOTAIR, we performed parallel experiments with the RAT-tagged WT and m6A mutant HOTAIR that were not incubated with METTL3/14. Thus, these transcripts would not be m6A methylated. Likewise, the m6A mutant HOTAIR transcript would not be m6A methylated because the methyltransferase complex is unable to establish the post-transcriptional modification at Uracil residues. These experiments will enable the comparison of YTHDCI binding in WT methylated HOTAIR, WT unmethylated HOTAIR, m6A mutant methylated HOTAIR, and m6A mutant unmethylated HOTAIR.

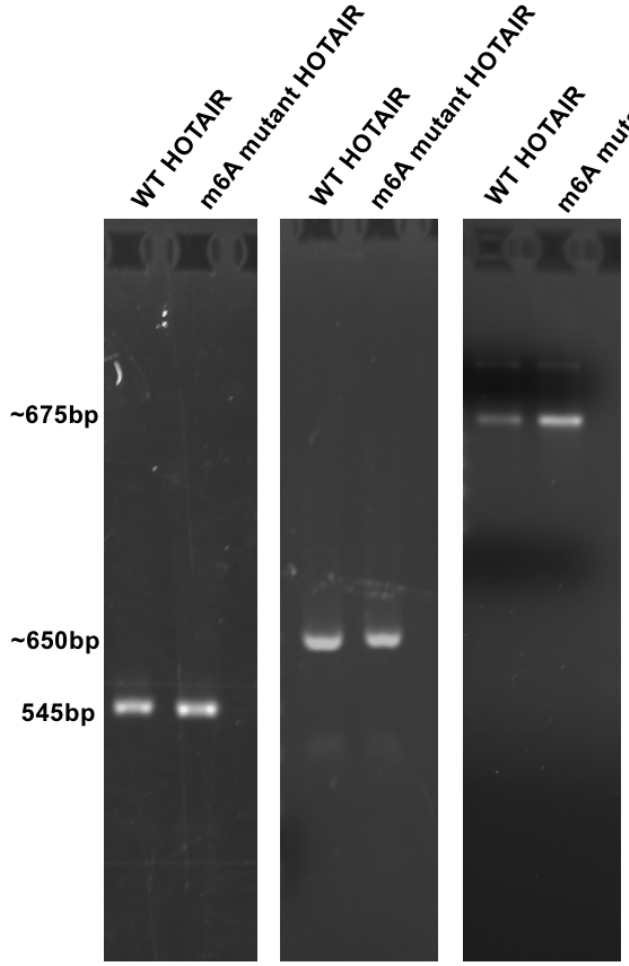

Figure 4. Successful IVT of WT and m6A-mutant domain 2 of HOTAIR. (Left) Partial templates for *in vitro* transcription were generated with PCR using Phusion polymerase and primers that had tails encoding a partial T7 promoter and the RAT tag and pLPCX and pBabePuro plasmids containing the m6A site of HOTAIR. (Middle) Full length DNA templates were created with PCR using Phusion polymerase and the purified partial template DNA from the initial round of PCR as well as the full length T7 promoter and RAT tag primers. (Right) WT and m6A mutant HOTAIR were *in vitro* transcribed using the MEGAscript T7 transcription kit. RNA Data shown is representative of one experiment.

YTHDC1 preferentially interacts with m6A HOTAIR

Because the ability of m6A to influence gene expression can involve the recruitment of specific reader proteins, we wished to determine whether the interaction between HOTAIR and YTHDC1 is dependent on the post-transcriptional modification [46-47]. To assay the recruitment of YTHDC1 to the m6A residue on HOTAIR, we used an RNA Affinity in Tandem (RAT) tag consisting of two hairpin stem-loops that selectively bind the *Pseudomonas aeruginosa* phage 7 (PP7). To determine the effects of the m6A modification on the HOTAIR-YTHDC1 interaction, we used our WT m6A methylated HOTAIR, wildtype m6A unmethylated HOTAIR, m6A mutant methylated HOTAIR, and m6A mutant unmethylated HOTAIR *in vitro transcripts*. Following the

binding of the RAT-tagged RNA to PP7, the samples were conjugated to IgG coupled dynabeads via the fusion protein A. Because we were unable to elute YTHDC1 but wanted to maintain a functional YTHDC1 protein, we incubated the RNA conjugated dynabeads with the unbound lysate from our anti-FLAG M2 affinity purification (Figure 3A). While the unbound lysate contained other nonspecific proteins, we reasoned that the amount of YTHDC1 present in the sample would be sufficient to perform a successful pulldown experiment. Indeed, the results from our RNA pulldown assays strongly suggest that YTHDC1 preferentially binds WT m6A methylated HOTAIR (Figure 5A). Using the statistical function of the Bio-Rad imager we calculated an approximate 1.7-fold enrichment of YTHDC1 for the methylated WT HOTAIR compared to the unmethylated WT HOTAIR (Figure 5B). Surprisingly, we calculated a 1.2-fold enrichment of YTHDC1 for methylated m6A mutant HOTAIR compared to unmethylated WT HOTAIR. Collectively, the results from these experiments establish a connection between the m6A pathway and HOTAIR via YTHDC1.

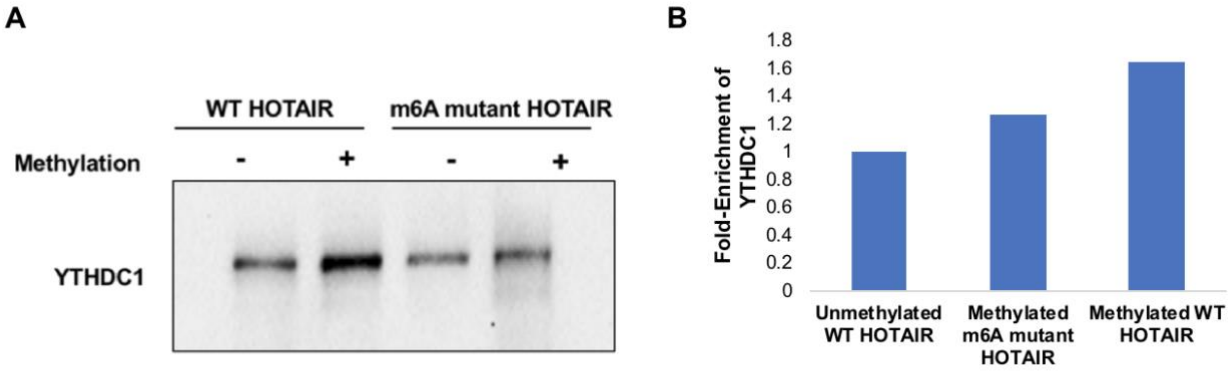

Figure 5. YTHDC1 enriched in WT methylated HOTAIR. (A) RAT-tagged RNA pulldown assay with WT unmethylated HOTAIR, WT methylated HOTAIR, m6A mutant unmethylated HOTAIR, and m6A mutant methylated HOTAIR. Pulldowns were analyzed with western blot analysis using an antibody against YTHDC1. (B) Fold-enrichment of YTHDC1 in WT m6A methylated and m6A mutant methylated HOTAIR were measured using the statistical analysis intensity bands obtained from the colorimetric image obtained in part A. Numbers were normalized with WT unmethylated HOTAIR.

DISCUSSION

Over the past decade, the field of ‘epitranscriptomics’ has established close to 200 hundred different chemical modifications to RNA that contribute to their overall biological function [25, 31, 33-48]. These studies have also identified many writers, erasers, and readers of these epigenetic modifications; however, their functional significance is only now being elucidated. The most common eukaryotic RNA modification, N6-methyladenosine (m6A), has emerged as a crucial regulator of normal development and metabolism [25, 36, 39-45, 48]. Notably, recent studies have shown that m6A has the unique capacity to recruit specific reader proteins to exert its downstream effects on gene expression [46-47]. While the ability of m6A to regulate RNA-protein interactions and RNA secondary structure has been extensively studied in mRNAs, the role of the post-transcriptional modification in promoting lncRNA function remains elusive. However, recent advancements in m6A detection methodology have identified many m6A modified lncRNAs. LncRNAs represent an exciting class of noncoding RNA transcripts that have been implicated in the regulation of gene expression at the epigenetic, transcriptional, and post-transcriptional levels of control [1-3, 6-15, 18-20]. The role of lncRNAs in regulating gene expression emphasizes the importance of exploring the role of the m6A modification on individual lncRNAs. In this present study, we focus on the lncRNA HOTAIR that acts *in trans* to epigenetically silence a 40-kb region of the *HOXD* gene cluster.

In this study, we screened HOTAIR for the presence of the m6A modification using the novel m6A eCLIP (meCLIP) method. This method allows for the detection of the m6A modification to single nucleotides by combining the eCLIP and m6A iCLIP (miCLIP). Because the aberrant expression of the m6A pathway and HOTAIR have been shown to promote breast cancer metastasis, we performed these studies in MCF-7 (ER+, PR+, HER2-) and MDA-MB-231 (triple-negative) breast cancer cells that express endogenously elevated levels of the lncRNA [9-13, 21-22, 25, 33, 36, 42, 44-46]. The results obtained from the meCLIP analysis were further validated with RNA immunoprecipitation experiments using an antibody against m6A. Our data

identified the m6A modification at nucleotide A783 of HOTAIR for both breast cancer cells. In addition, the results obtained from these experiments also detected the m6A modification on many of HOTAIR's mRNA targets. The fact that HOTAIR and many of its target mRNAs are m6A modified suggests that the post-transcriptional modification has some biological significance.

Recently, the lncRNA XIST was shown to contain multiple m6A residues indicating that the post-transcriptional modification is biologically significant [44]. Indeed, the m6A residues on XIST have a demonstrated ability to recruit YTHDC1 to promote the transcriptional silencing of one X-chromosome during early embryonic development in female mammals. YTHDC1, a member of a family of proteins known to recognize m6A residues, is proposed to regulate numerous RNA processing events. The discovery that m6A recruits specific reader proteins solidified the notion that post-transcriptional modifications demonstrate regulatory potential. Because recent improvements in screening methods have identified many additional m6A modified lncRNAs, it is likely that the post-transcriptional modification contributes to many diverse pathophysiological mechanisms.

Several aspects of the mechanism underlying HOTAIR-mediated gene repression are not entirely clear, including how HOTAIR is recruited to specific sites in the genome to trigger heterochromatin formation and the methyltransferase activity of PRC2. Previously, HOTAIR was proposed to act as a transcriptional repressor through the recruitment of PRC2, which deposits H3K27me3 marks at specific sites in the genome [2-3, 7-9, 12-15]. However, the specificity of PRC2 was questioned by the finding that the methyltransferase complex readily binds multiple RNAs on chromatin [28]. In addition, HOTAIR was shown to maintain its repressive function in the absence of PRC2 through an unknown mechanism [29]. Taken together, these results strongly suggest that additional factors contribute to HOTAIR-mediated gene repression and

that the lncRNA itself may play an important role in the observed specificity. Given the regulatory role of m6A in XIST-mediated gene repression, it is tempting to speculate the m6A residue of HOTAIR contributes to the observed gene repression mechanism by mediating interactions with proteins. Because HOTAIR-mediated gene repression remains relatively unclear, the finding that HOTAIR contains the m6A modification proposes an exciting potential new mechanism of action for the lncRNA.

To explore whether the m6A modification on HOTAIR mediates interactions with proteins, we performed RNA immunoprecipitation experiments. We focused are efforts on YTHDC1 because of its known role in XIST-mediated transcriptional repression and other RNA processing events. Because the overexpression of HOTAIR significantly alters gene expression profiles to promote cancer pathogenesis, we performed these experiments in MCF-7 and MDA-MB-231 breast cancer cells that express endogenously elevated levels of the lncRNA [2-3, 7-9, 12-15, 22]. The results from these experiments indicate that HOTAIR and many of its target mRNAs associate with YTHDC1, however, they do not address the molecular details of the interactions. Although these experiments do not establish a concrete connection between the m6A pathway and HOTAIR-mediated gene repression, they present a potential new mechanism of action for the lncRNA.

The fact that HOTAIR and many of its mRNA targets are m6A modified and interact with YTHDC1 suggests that the post-transcriptional modification may play an important role in the observed gene repression mechanism. The m6A modification has been shown to mediate RNA-protein interactions to influence RNA processing events [23, 31, 33-45]. In light of these findings, we sought to explore the HOTAIR-YTHDC1 interaction and, in particular, whether it is dependent on the m6A modification. To do this, we performed RNA pulldown experiments with *in vitro* transcribed wild-type and m6A mutant HOTAIR in the presence or absence of

methylation. Our data reveals that YTHDC1 preferentially interactions with WT m6A methylated HOTAIR, suggesting that the post-transcriptional modification mediates the HOTAIR-YTHDC1 interaction. However, we also observed an enrichment of YTHDC1 associating with m6A mutant methylated HOTAIR. This unexpected data may be explained by the impurity of the unbound lysate, which contained nonspecific proteins that may have interfered with the HOTAIR-YTHDC1 interaction. However, it is probably more likely that our *in vitro* methylation experiments lacked specificity. For instance, the combination of METTL3, METTL14, and SAM may have methylated additional unexpected sites in the transcript, raising the concern of off target methylation interfering with the data from the study. If another site was methylated in the m6A mutant HOTAIR, it is possible that it was able to bind YTHDC1. Future studies performing meCLIP on *in vitro* modified HOTAIR will determine whether the methods we used to methylate transcripts are successful in establishing individual m6A residues. Although additional studies are needed to further elucidate the biological significance of this interaction, our findings offer a promising new avenue of research that could lead to the better understanding of HOTAIR-mediated gene repression. It will be important to determine whether the methylation of the adenosine residue directly contributes to HOTAIR-induced gene silencing.

Although the detailed mechanisms by which m6A exerts its effects on gene repression are not entirely understood, it is becoming increasingly evident that there is significant cross-talk between RNA and histone modifications. For instance, the m6A RNA modification has been shown to regulate histone modifications like H3K27me3 [40]. Given the presence of the m6A modification within the HOTAIR transcript and many of its target genes, as well as the H3K27me3 repressive mark, it is possible that the two mechanisms interact to regulate gene expression. Indeed, m6A alters the stability of the mRNAs that encode histone modifiers to ultimately regulate histone modifications. To test this, we could perform knockout experiments of the m6A methyltransferase complex components including, METTL3 and METLL14. Following

the successful knockout, we could test for levels of specific histone modifications like H3K27me3. If there is an overlap between the two pathways, then we would expect to find decreased levels of histone modifications following the knockout of METTL3 and METLL14.

The identification of the m6A modification in the lncRNA may have important implications in the development of therapeutics for metastatic breast cancer. Perhaps one of the most exciting aspects of m6A is its potential to be reversed by its writer proteins. Because FTO is frequently downregulated in invasive breast cancer and the overexpression of HOTAIR is linked to increased metastatic ability, it is tempting to speculate that the two pathways are connected [44]. While beyond the scope of this present study, future studies investigating how the HOTAIR-YTHDC1 interaction contributes to metastatic breast cancers will likely lead to a better understanding of the observed gene repression. This is extremely important because breast cancer represents a highly heterogeneous pathology that is a leading cause of death for women worldwide. The American Cancer Society estimates that there were over 63,000 cases of *in situ* breast cancer diagnosed in 2017, and approximately 40,000 deaths stemming from invasive breast cancer [48]. Thus, breast cancer remains a significant global healthcare problem despite increased screening availability and advancements in effective treatment options. Ultimately, by identifying the proteins that interact with HOTAIR to induce gene repression in aggressive cancers, we will gain a better understanding of how lncRNAs regulate physiological processes.

The long noncoding RNA HOTAIR presents a promising therapeutic avenue for aggressive breast carcinomas because of its critical role in regulating gene expression, which in turn controls cellular proliferation, migration, and invasion. For instance, the overexpression of HOTAIR in MDA-MB-231 breast cancer cells was shown to promote lung metastasis *in vivo* [13, 49]. Moreover, depleting HOTAIR with siRNA in MCF-7 breast cancer cells was sufficient to decrease invasion. Additionally, recent data has shown that the suppression of HOTAIR

expression induced apoptosis and sensitized cancer cells to tumor necrosis factor α (TNF- α) and to the chemotherapeutic agents' cisplatin and doxorubicin [49]. This finding suggests that HOTAIR could be used as an effective therapy for drug resistant malignancies. Because lncRNA expression is cancer- and tissue-specific, anti-tumor strategies that focus on lncRNAs will likely be capable of selectively targeting cancer cells [50]. Given that current therapeutic options are unable to specifically kill cancer cells, cancer-associated lncRNAs represent a powerful potential anti-cancer target. However, there are many obstacles that need to be overcome before feasible cancer therapeutics can be established. Epigenetic gene regulation is highly complex and dynamic. As each chromatin regulator may affect numerous genes throughout the genome, it will be extremely important to develop new methodology that allows for the detection of downstream effects of chromatin regulators. This would be invaluable tool in the development of HOTAIR therapeutics because it could predict the tissue-specific outcomes when the drug was administered. Although previous studies have been able to suppress HOTAIR expression in multiple different ways, including small molecule inhibitors, antagonistic oligonucleotides, and siRNAs, the off-target effects of these methods have yet to be explored [49]. Thus, additional research into the gene repression mechanism and associated molecules is needed to better understand the potential use of HOTAIR as an anti-cancer therapeutic.

WORKS CITED

1. ENCODE Project Consortium, Birney E, Stanatoyannopoulos J.A., Dutta A, Guigo R, Gingeras T.R., *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007. **447**:799-816.

2. Bhan A, Mandal S.S. LncRNA HOTAIR: a master regulator of chromatin dynamics and cancer. *Biochem Biophys Acta* 2015. **1856**: 151-164.

3. Pawlowska E, Szczezepanska J, Blasiak J. The Long Noncoding RNA HOTAIR in Breast Cancer: Does Autophagy Play a Role*? Int. J. Mol. Sci* 2017; **18**(11).

4. Liu W, Wu J, Shi G, Yue X, Liu D, Zhang Q. Aberrant promoter methylation of PCDH10 as a potential diagnostic and prognostic biomarker for patients with breast cancer. *Oncology Letters* 2018. **16**: 4462-4470.

5. Xing Z, Park P.K., Lin C, Yang L. LncRNA BCAR4 wires up signaling transduction in breast cancer. *RNA Biology* 2015. **12**(7): 681-689.

6. Castro-Oropeza R, Melendez-Zajgla J, Maldonado V, Vasquez-Santillan K. The emerging role of lncRNAs in the regulation of cancer stem cells. *Cellular Oncology* 2018. [epub ahead of print].

7. Balas M.M., Johnson A.M. Exploring the mechanisms behind long noncoding RNAs and cancer. *Non-coding RNA research* 2018. 1-10.

8. Kondo Y., Shinjo K., Katsushima K. Long non-coding RNAs as an epigenetic regulator in human cancers. *Cancer Sci* 2017. **108**: 1927-1933.

9. Meredith E.K., Balas M.M., Sindy K, Haislop K, Johnson A.M. An RNA matchmaker protein regulates the activity of the long noncoding RNA HOTAIR. *RNA* 2016. **22**: 1-16.

10. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, *et al*. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 2011; **477**(7364): 295-300.

11. Giles KE, Woolnough JL, Atwood B. 2015. ncRNA function in chromatin organization. In: Huang S, Litt MD, Blakey CA, editors. Epigenetic Gene Expression and Regulation. Academic Press. p. 117-148.

12. Schmitt, AM., Chang, HY. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell* 2016; 29(4): 452-463.

13. Gupta R.A., Shah N, Wang K.C., Kim J, Horlings H.M., Wong D.J., Tsai M, Hung T, Argani P, Rinn J.L., *et al*. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature Letters* 2010. **464**: 1071-1076.

14. Nguyen E.D., Balas M.M., Griffin A.M., Roberts J.T., Johnson A.M. Global profiling of hnRNP A2/B1-RNA binding on chromatin highlights LncRNA interactions. *RNA Biology* 2018. **15**(7): 901-913.

15. Rinn JL, Kertesz M, Wang J.K., Squazzo SL, Xu X, Brugmann SA, Goodnough H, Helms JA, Fernham PJ, Segal E, *et al.* Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Non-Coding RNAs. *Cell* 2007. **129**(7): 1311-1323.

16. Turner, B.M., 2001, Chromatin and gene regulation: mechanisms in epigenetics, The Americas Blackwell, Williston, VT, 283 p.

17. Vidal, M, 2009, Chromatin Modifications by Polycomb Complexes, in Esteller, M, ed., *Epigenetics in Biology and Medicine*, CRC Press, Boca Raton, p. 131-155.

18. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci* 2009. 106(28): 11667-11672.

19. Bannister, A.J., Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res*. 2011. 21(3): 381-395.

20. Han P, Chang CP. Long non-coding RNA and chromatin remodeling. *RNA Biol.* 2015. 12(10): 1094-1098.

21. Arab K, Park YJ, Lindroth AM, Grummt I, Niehrs C, Plass C. Long Noncoding RNA *TARID* Directs Demethylation and Activation of the Tumor Suppressor *TCF21* via DADD45A. *Mol. Cell*. 2014. 55(4): 604-614.

22. Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. *Nat Rev Cancer*. 2018. 18(1): 5-18.

23. Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP, Cui H. Epigenetic silencing of tumor suppressor gene *p15* by its antisense RNA. *Nature*. 2008. 451(7175): 202-206.

24. Wolffe, A, 1998, Chromatin structure & function, Academic Press, San Diego, CA, 447 p.

25. Liu Z, Li L, Sun H, Liu S. Link Between m6A Modification and Cancers. Front. *Bioeng. Biotechnol.* 2018. **6**: 89.

26. Tang Q, Hann S.S. HOTAIR: An Oncogenic Long Non-Coding RNA in Human Cancer. *Cell Physiol Biochem* 2018. **47**: 893-913.

27. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010. **329**:689-693.

28. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat Struct Mol Bio* 2013. **20**:1250-1257.

29. Portoso M, Ragazzini R, Brenčič Ž, Moiani A, Michaud A, Vassilev I, Wassef M, Servant N, Sargueil B, Margueron R. PRC2 is dispensable for *HOTAIR*-mediated transcriptional repression. *The EMBO Journal* 2017. **36**(8): 981-994.

30. Gao S, Zhou B, Li H, Huang X, Wu Y, Xing C, Yu X, Ji Y. Long noncoding RNA HOTAIR promotes the self-renewal of leukemia stem cells through epigenetic silencing of p15. *Experimental Hematology* 2018. **00**:00-00.

31. Ropero S, Esteller M, 2009, Epigenetics and Cancer: DNA Methylation, in Esteller, M, ed., *Epigenetics in Biology and Medicine*, CRC Press, Boca Raton, p. 3-17.

32. Cheng D, Deng J, Zhang B, He X, Meng Z, Li G, Ye H, Zheng S, Wei L, Deng X. LncRNA HOTAIR epigenetically suppresses miR-122 expression in hepatocellular carcinoma via DNA methylation. *E Bio Medicine* 2018. **34**: 76-84.

33. Smith, M, 2011, Investigating the Human Genome: Insights into Human Variation and Disease Susceptibility, FT Press, Upper Saddle River, NJ, 216 p.

34. Elliot, D, Ladonmery, M, Molecular Biology of RNA, Oxford University Press, New York City, 441 p.

35. Allis, DC., Jenuwein, T, Reinberg, D., 2007, Epigenetics, CSHL Press, Cold Spring Harbor, NY, 502 p.

36. Fitzsimmons, CM, Batista PJ. It's complicated… m6A-dependent regulation of gene expression in cancer. *BBA-Gene Regulatory Mechanisms* 2018. https://doi.org/10.1016/j.bbagrm.2018.09.010.

37. Mizrahi A, Czerniak A, Levy T, Amiur S, Gallula J, Matouk I, *et al*. Development of targeted therapy for ovarian cancer mediated by a plasmid expression diphtheria toxin under the control of H19 regulatory sequences. *J Transl Med* 2009; **7**:69.

38. Yang D, Qiao J, Wang G, Lan Y, Li G, Guo X, Xi J, Ye D, Zhu S, Chen W, et al. N^6-Methyladenosine modification of lincRNA 1281 is critically required for mESC differentiation potential. *Nucleic Acids Research* 2018. **46**: 3906-3920.

39. Gilbert W.V., Bell T.A., Schaening C. Messenger RNA modifications: Form, distribution, and function. *Science* 2016. **352**: 1408-1412.

40. Meyer K.D., Jaffrey S.R. The dynamic epitranscriptome: N^6-methyladenosine and gene expression and control. *Nature reviews* 2014. **15**: 313-326.

41. Tong J, Flavell R.A., Li H. RNA m^6A modification and its function in disease. *Front. Med* 2018. **12**(4): 481-489.

42. Liao S, Sun H, Xu C. YTH Domain: A Family of N^6-methyladenosine (m^6A) Readers. *Genomics Proteomics Bioinformatics* 2018. **16**: 99-107.

43. Xiao W, Adhikari S, Dahal U, Wang X, Zhao Y, Yang Y. Nuclear m^6A Reader YTHDC1 Regulates mRNA Splicing. *Molecular Cell* 2016. **61**: 507-519.

44. Kaklamani V, Yi N, Sadim M, Siziopikou K, Zhang K, Xu Y, et al. The role of the fat mass and obesity associated gene (*FTO*) in breast cancer risk. *BMC Medical Genetics* 2011, **12**:52.

45. Hong, K. Emerging function of N6-methyladenosine in cancer. *Oncol Lett.* 2018. 16(5): 5519-5524.

46. Grosjean H. Modification and editing of RNA: Historical overview and important facts to remember. In Grosjean H. Fine-Tuning of RNA Functions by Modification and Editing. Topics in Current Genetics, Volume 12. Springer, Berlin, Heidelberg.

47. Patil D.P., Chen C.K., Pickering B.F., Chow A, Jackson C, Guttman M, Jaffrey S.R. m^6A RNA methylation promotes *XIST*-mediated transcriptional repression. *Nature* 2016. **537**: 369-373.

48. Wang Y, Li Y, Yue M, Wang J, Kumar S, Wechsler-Reya RJ, Zhang Z, Ogawa Y, Kellis M, Duester G, Zhao JC. N6-methyladenosine RNA modification regulates embryonic

neural stem cell self-renewal through histone modifications. *Nature Neuroscience* 2018. **21**: 195-206.

49. DeSantis CE, Ma J, Sauer AG, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin* 2017. 67(6): 439-448.

50. Yao Y, Li J, Wang L. Large Intervening Non-Coding RNA *HOTAIR* is an Indicator of Poor Prognosis and a Therapeutic Target in Human Cancers. *Int J Mol Sci*. 2014. 15(10): 18985-18999.

ACKNOWLEDGEMENTS

This work was supported by the RNA Bioscience Initiative Summer Internship Program at the University of Colorado School of Medicine. I would like to express my gratitude to Dr. Aaron Johnson and Dr. Allison Porman for their continuous support, patience, and guidance. I also want to thank Justin Roberts, Stephen Wu, Dr. Alexis Zukowski, Dr. Maggie Balas, April Griffin, and Chloe Barrington-Ham for being the best lab mates I could have asked for.

I would also like to thank Dr. Darrell Killian and Dr. Jennifer Garcia at Colorado College for helping me with my thesis. I am so grateful to have had your continuous encouragement and feedback throughout the entire process.

Finally, I want to thank the rest of the Molecular Biology professors for continually challenging me to think deeper and for inspiring me to pursue a career in science. MB has been an integral part of my past four years and I am so grateful to have learned from them.

CALCULATION APPENDIX

RNA	Band No.	Volume (Int.)	Rel. Quant.	Fold Enrichment of YTHDC1*
Lane 1 *Methylated m6A mutant HOTAIR*	1	3869748	1	**1.264817335**
	2	687132	0.177565	
	3	11490588	2.969338	
	4	3753738	0.970021	
	5	82872	0.021415	
	6	169056	0.043687	
Lane 2	1	513666	0.132739	

Unmethylated m6A mutant HOTAIR				
	2	3059532	0.790628	
	3	9354456	2.41733	
	4	5342652	1.38062	
	5	64800	0.016745	
	6	231750	0.059888	
Lane 3 *Methylated WT HOTAIR*	1	255060	0.065911	**1.652773973**
	2	5976252	1.544352	
	3	666432	0.172216	
	4	10149930	2.622892	
	5	5282298	1.365024	
Lane 4 *Unmethylated WT HOTAIR*	1	3615894	0.9344	
	2	11253438	2.908054	
	3	5802750	1.499516	
	4	426348	0.110175	

*Statistical analysis was performed using the Bio-Rad imager software. Data was normalized to YTHDC1 pulldown with unmethylated WT HOTAIR. Bands were selected for relative intensity and quantities for each of the bands present for the four samples.