

AWARD NUMBER: W81XWH-10-1-0567

TITLE: A long stress-responsive non-coding transcript (Nt 5) and its role in the development of breast cancer

PRINCIPAL INVESTIGATOR: David I Smith, Ph.D.

CONTRACTING ORGANIZATION: Mayo Clinic
Rochester, MN 55905

REPORT DATE: July 2011

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.				
1. REPORT DATE (DD-MM-YYYY) 1 August 2011		2. REPORT TYPE Annual		3. DATES COVERED (From - To) 8/1/2011-7/31/2011
4. TITLE AND SUBTITLE A long stress-responsive non-coding transcript (Nt 5) and its role in the development of breast cancer		5a. CONTRACT NUMBER W81XWH-10-1-0567		
		5b. GRANT NUMBER		
		5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) David I Smith, P.hD. smith.david@mayo.edu		5d. PROJECT NUMBER		
		5e. TASK NUMBER		
		5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Mayo Clinic Rochester, MN 55905		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research & Materiel Command Fort Detrick, MD 21702-5012		10. SPONSOR/MONITOR'S ACRONYM(S)		
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT We received funding from the Department of Defense Breast Cancer Program to study a long stress-responsive non-coding transcript which is called LSINCT5. There were three Specific Aims to our research proposal whose overall goal was to better understand the structure and function of this long stress-responsive non-coding transcript which was found to be frequently over-expressed in breast cancer as compared to matched normal breast epithelial tissue. Our three Specific Aims were to (1) Characterize the LSINCT5 transcript to determine its full length, which RNA polymerase transcribes it, and its sub-cellular localization; (2) To analyze the phenotype effect of modulating the expression of LSINCT5 in a normal breast epithelial cell line and in breast cancer cell lines that over-express it; and (3) To determine what other genes and non-coding transcripts have altered expression in direct response to altering the expression of LSINCT5. In this annual progress report we summarize our work on the analysis of LSINCT5, and we have completed most of the proposed work that we originally outlined. However, the most important goal is to determine what is the function of LSINCT5 in normal cells and how its over-expression in breast cancers contributes to breast cancer development. We have, therefore, initiated experiments aimed at determining what are the genes and non-coding transcripts that interact with LSINCT5 so that we can unravel the precise cellular functions of this interesting non-coding transcript.				
15. SUBJECT TERMS non-coding transcript, stress-responsive. Breast cancer proliferation, molecular genetics, cellular regulators				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 23
a. REPORT	b. ABSTRACT U	c. THIS PAGE U		
			19a. NAME OF RESPONSIBLE PERSON	
			19b. TELEPHONE NUMBER (include area code)	

Table of Contents

	Page
Abstract.....	3
Introduction.....	4
Body.....	4
Key Research Accomplishments	12
Publications as a Direct Result of this Grant	12
Abstracts as a Direct Result of this Grant	12
Presentations Resulting from this Grant	12

ABSTRACT

Background: Most of the focus in breast cancer genetics has been the study of specific genes whose alteration either predisposes or is involved in breast cancer development. However, less than 2% of the human genome actually is involved in coding for protein. It has recently become clear that much of the non-coding portion of the genome is still transcriptionally active and that non-coding transcripts vastly exceed coding transcripts. The non-coding transcripts have been found to play important regulatory roles within cells including breast cancer cells. We have now identified a group of long stress-responsive non-coding transcripts that appear to be involved in promoting cellular proliferation in breast cancer.

Objective: We propose to characterize one of the newly identified stress responsive long non-coding transcript as this transcript is derived from a chromosomal region frequently amplified in breast cancers and it is over-expressed in most breast cancers cell lines and primary tumors tested to date. Our hypothesis is that this non-coding transcript is representative of a new group of transcripts that promote proliferation and as such are important targets during breast tumorigenesis. We will characterize this transcript to determine the role that it plays in the development of breast cancer and determine the genes and pathways that it is involved in regulating. This and other long stress-responsive non-coding transcripts might be novel therapeutic targets for the treatment of breast cancer.

Specific Aims: There are three specific aims to this proposal. The first specific aim is to fully characterize this transcript including determining where it begins and ends, its localization within both normal breast epithelium and in breast cancers, and to identify where and when it is over-expressed. The second specific aim is to knock-down the expression of the over-expressed transcript in breast cancer cell lines and also to over-express it in normal breast epithelial cell lines. We will then observe the resulting phenotypic effect of modulating the expression of this transcript. The third specific aim is to characterize the effect of this modulation on the entire transcriptome using Next Generation DNA sequencing on the SOLID/ABI platform.

Innovation: This study is examining a newly identified stress-responsive non-coding transcript as a representative of a new group of regulatory transcripts that play an important role in the development of breast cancer. Little work has been done characterizing the role that long non-coding transcripts play in breast cancer development. Several of the considerably shorter microRNAs have been found to play important regulatory roles and to be important targets in breast tumorigenesis. There have been recent discoveries suggesting that the role of non-coding transcript is to regulate coding gene expression. We believe that our stress-responsive long non-coding transcript promotes cellular proliferation and is just one of possibly many long non-coding transcripts that could be involved in breast cancer development. After modulating its expression and observing the resulting phenotypic changes that it has on normal breast epithelial cells, we will use the power of Next Generation DNA sequencing to examine the effect it has on the entire transcriptome of the breast cells, not just on the coding portion of the genome.

Impact: The identification and characterization of regulatory small microRNAs has revealed that several miRNAs are important targets of alteration in the development of breast cancer. We believe that the longer stress-responsive non-coding transcripts are also important cellular regulators that play an important role in breast tumorigenesis. They may also be important therapeutic targets. Hence, we believe that these studies will reveal new transcripts that play important roles in breast cancer development.

INTRODUCTION

BODY

Original Proposal

There were three Specific Aims to our proposal. They were:

Specific Aim #1: Characterization of LSINCT5

We will first more fully characterize LSINCT5 (formally called NIT5). We will determine where this transcript begins and ends using 5' and 3' RACE as well as RNase protection assays. We will also characterize which polymerase is involved in its transcription. Last, we will identify where the LSINCT5 transcript resides in both normal and breast cancer cells.

Specific Aim #2: Analyze the phenotypic effect of modulating LSINCT5 expression

We will use siRNA technology to decrease the expression of LSINCT5 in breast cancer cell lines where it is most highly over-expressed. We will also construct a full-length LSINCT5 expression construct to over-express it in normal human breast epithelial cell lines (HMEC and MCF10A). We will also make stable constructs with either increased or decreased expression of LSINCT5. We will then use a variety of different assays including proliferation assays (MTT and BrDU stain), migration assay (Boyden chambers), soft agar assay, and apoptosis assays (Annexin V/PI stain, LIVE/DEAD Fixable Dead Cell stain Kit, and Tunnel Assay) to determine the phenotypic effect of modulating LSINCT5 expression.

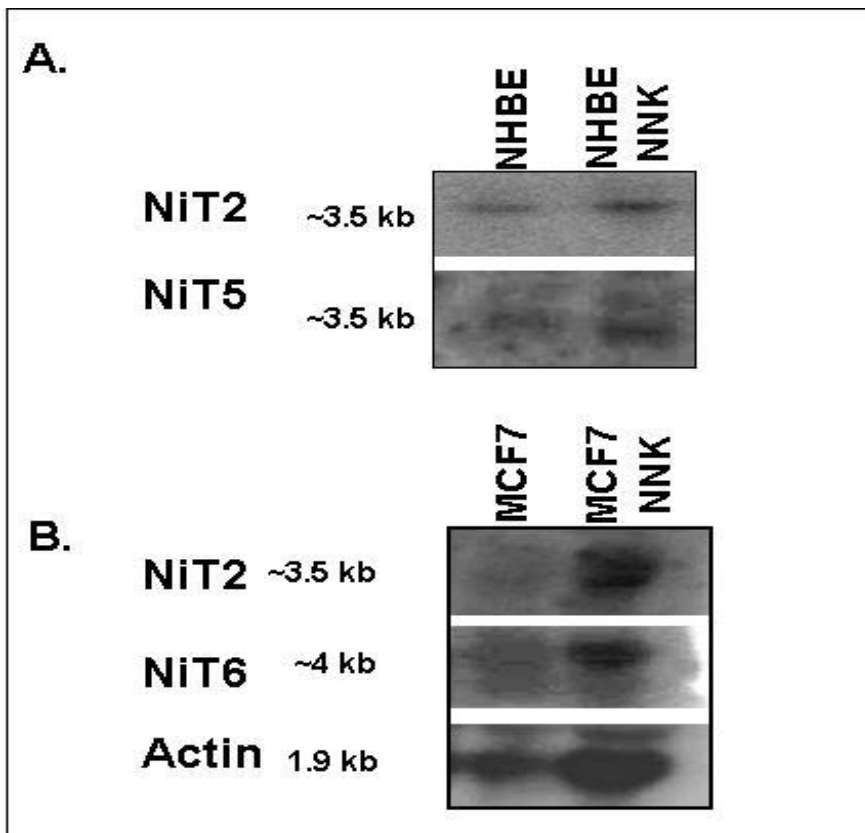
Specific Aim #3: To characterize genes, non-coding transcripts and pathways that are affected by the modulation of LSINCT5 expression

We will utilize the power of Next Generation DNA sequencing to characterize the effect of modulating the expression of LSINCT5 on the entire transcriptome of either normal breast epithelial cell lines or breast cancer cell lines. Hence, we will be able to characterize which coding genes are affected by modulating LSINCT5 expression as well as non-coding transcripts that have altered expression as a result of changing LSINCT5 expression. This work will enable us to determine how LSINCT5 is involved in cellular proliferation and which genes and pathways LSINCT5 interacts with.

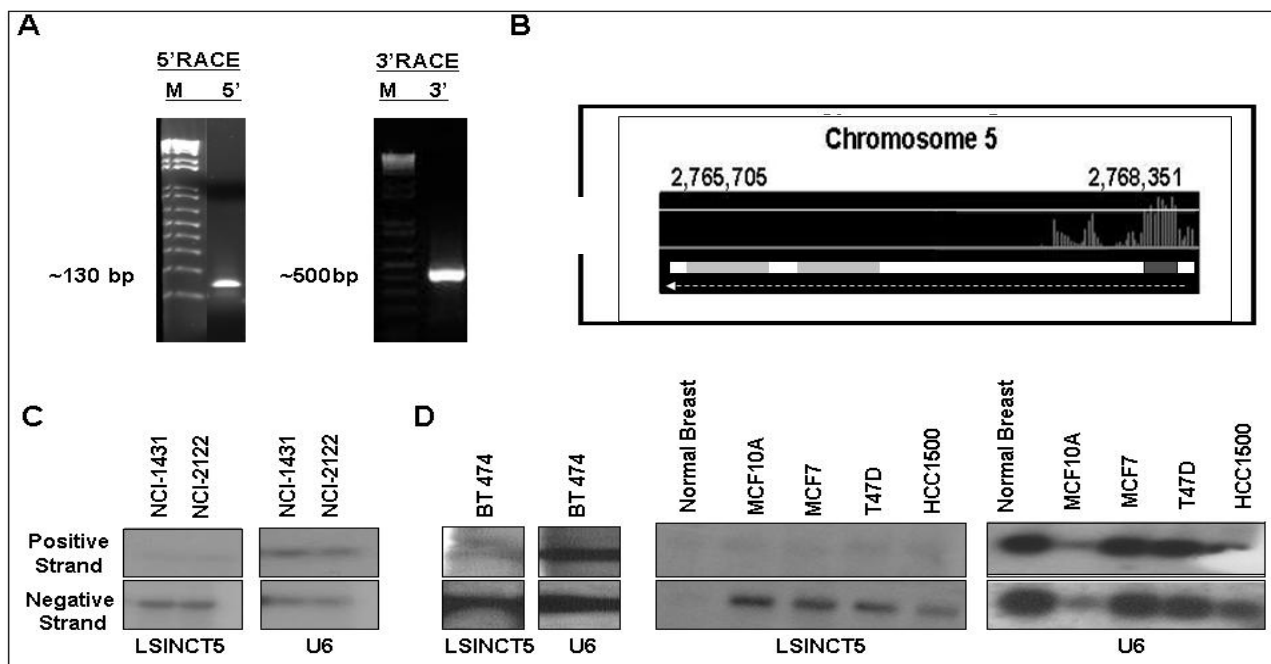
Specific Aim #1: Characterization of LSINCT5

We will first more fully characterize LSINCT5 (formally called NIT5). We will determine where this transcript begins and ends using 5' and 3' RACE as well as RNase protection assays. We will also characterize which polymerase is involved in its transcription. Last, we will identify where the LSINCT5 transcript resides in both normal and breast cancer cells.

We started our work on the characterization of LSINCT5 by first utilizing 5' and 3' RACE to determine where this long non-coding transcript begins and ends. The “core” sequence of LSINCT5 is that 350 base sequence which was first detected using the tiling array on NHBE cells both before and after treatment with the tobacco carcinogen NNK. Overlapping oligonucleotides within that core sequence hybridized more strongly to RNA isolated from NHBE cells treated with NNK than to control NHBE cells. However, when we took that “core” sequence and utilized it as a probe against Northern blots produced from NHBE cells treated with NNK, we found that the apparent full length transcript of LSINCT5 was almost 3,000 bases in length. Indeed many of the 12 NNK-induced transcripts that we first identified (Silva et al., 2010) hybridized to transcripts that were considerably larger than the apparent size of the transcripts as detected from the tiling array experiment. It is important to realize that the construction of the tiling array was based upon making probes across the non-redundant portion of the genome; there were no oligos constructed across repetitive elements such as SINEs or LINEs.

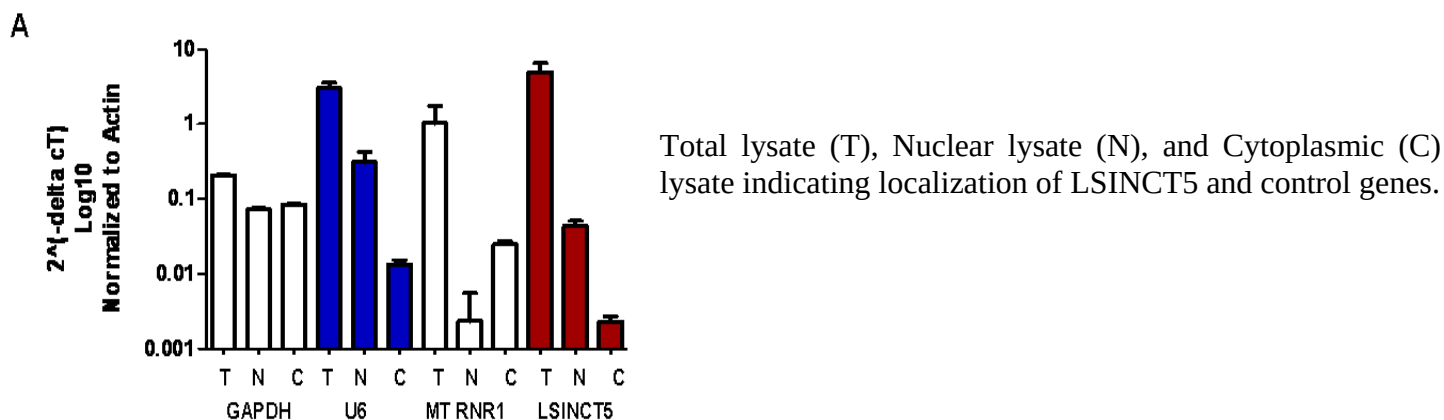


We first did 3' RACE to isolate the 3' end of the full length LSINCT5 transcript. The 3' RACE with a probe derived from within the "core" LSINCT5 sequence yielded a 500 base pair fragment; the 3' end was relatively close to the end of the "core" sequence. Since the full length LSINCT5 transcript was almost 3 Kb in length, we expected the 5' end of the transcript to lie much further upstream of the "core" sequence. We constructed several oligonucleotides based upon sequences upstream of the "core" sequence and one oligonucleotide primer derived over 1 Kb upstream of the "core" sequence yielded a 130 bp fragment with the 5' RACE. Hence the full length of the LSINCT5 transcript is 2,647 bases in size. The 5' end of the LSINCT5 transcript contains 2 LINE repeat elements which is the reason why there were not many oligonucleotide probes on the tiling array covering that portion of LSINCT5. The Figure below shows the 5' and 3' RACE fragments rescued as well as the tiling array data for the full length LSINCT5 transcribed region. We also cloned the LSINCT5 "core" sequence into a directional cloning plasmid so that we could produce both a positive and negative strand hybridizing probe. Only the negative strand probe hybridized to the LSINCT5 transcript, and so LSINCT5 is transcribed from the negative strand; this is shown in part C of the Figure below. Part D shows the hybridization of the positive and negative strand probes to various breast cancer cell lines (as described in the legend under the Figure).



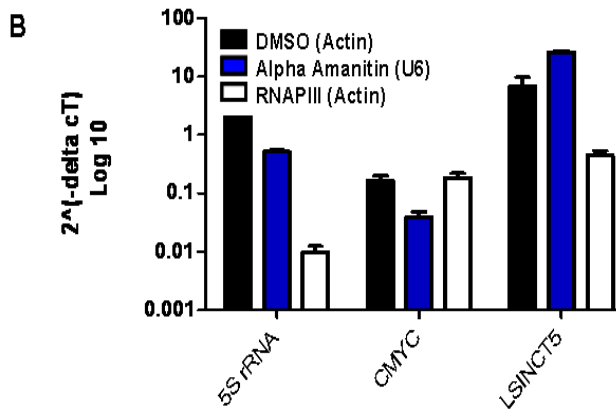
A: 5' RACE and 3' RACE indicating beginning and end of LSINCT5. **B:** Tiling array data analyzed with IGB for LSINCT5 and chromosomal characteristics including 2 LINE repeat elements (light grey bars) and core sequence (black bar) in NHBE-NNK-induced cells. **C:** Directional Northern Blot showing LSINCT5 expression in negative strand in two lung cancer cell lines and breast cancer panel

To determine where within cells LSINCT5 is localized we isolated a nuclear and cytoplasmic fraction from HMEC cells as well as from MCF10A. We then used real-time RT-PCR to measure expression of LSINCT5 as well as GAPDH (whose transcript is detected in both the nuclear and cytoplasmic fraction), a control nuclear gene (U6) and a control cytoplasmic gene (MT RNR1). We detected more of the U6 transcript in the nuclear than the cytoplasmic fraction and conversely more of MT RNR1 in the nuclear fraction. With the LSINCT5 probe we found that it had considerably more signal in the nuclear than the cytoplasmic fraction, hence LSINCT5 is most probably a nuclear long non-coding transcript. The results from this experiment are shown in the Figure below.



Total lysate (T), Nuclear lysate (N), and Cytoplasmic (C) lysate indicating localization of LSINCT5 and control genes.

We next wanted to determine which polymerase was responsible for LSINCT5 transcription. To determine this we utilized an inhibitor of RNA polymerase II (alpha amanitin) and an inhibitor of RNA polymerase III (RNAPIII). Our control genes in this experiment were 5S rRNA (which is transcribed by RNA polymerase III) and c-myc (which is transcribed by RNA polymerase II). The figure below shows the results from this experiment which clearly indicate that LSINCT5 transcription is only inhibited by the RNA polymerase III inhibition.



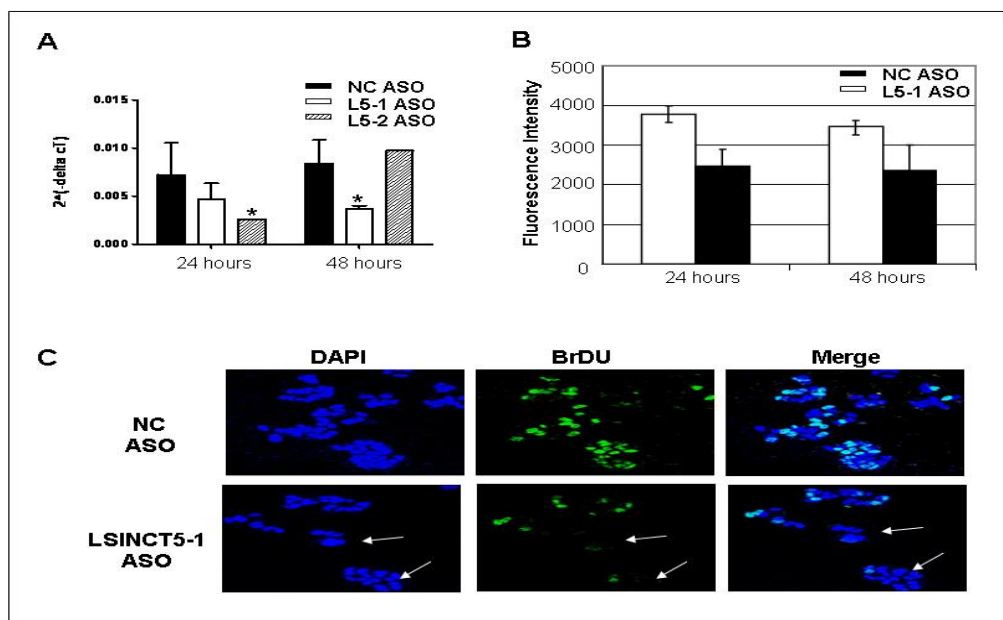
RNA Polymerase II (Alpha amanitin) and RNA Polymerase III (RNAPIII) treatments showing LSINCT5 expression and control genes. qPCR experiments conducted in triplicate and repeated two times.

We were therefore able to successfully complete each of our goals for the first Specific Aim. There was one additional thing that we wanted to do for this Specific Aim and that was to determine the exact location of LSINCT5 both in normal and breast cancer derived cell lines. Those results have been obtained and will be described under the third Specific Aim whose major goal was to characterize genes whose expression was altered when we altered LSINCT5 expression.

Specific Aim #2: Analyze the phenotypic effect of modulating LINCT5 expression.

We will use siRNA technology to decrease the expression of LSINCT5 in those breast cancer cell lines where it is most highly over-expressed. We will also construct a full-length LSINCT5 expression construct to over-express it in normal human breast epithelial cell lines (HMEC and MCF10A). We will also make stable constructs with either increased or decreased expression of LSINCT5. We will then use a variety of different assays including proliferation assays (MTT and BrDU stain), migration assay (Boyden chambers), soft agar assay, and apoptosis assays (Annexin V/PI stain, LIVE/DEAD Fixable Dead Cell stain Kit, and Tunnel Assay) to determine the phenotypic effect of modulating LSINCT5 expression.

We started this Specific Aim trying to use siRNAs specific for LSINCT5 to modulate its expression. Unfortunately, none of the synthesized siRNAs were able to decrease LSINCT5 mRNA expression. However, LSINCT5 is a nuclear transcript so we did a literature search looking for another strategy that might prove more successful with transcripts that are confined to the nucleus; we reasoned that the siRNA was not working because it was unable to get into the nucleus to decrease LSINCT5 transcripts levels. The use of antisense oligonucleotides (ASOs) has recently been shown to successfully knock down the expression of nuclear nc RNAs (Hamilton TC. *Semin Oncol* 1984; 11:285-298). We created two ASOs specific to LSINCT5 (LSINCT5-1 and LSINCT5-2 ASOs) and a scrambled normal control (NC ASO) to knock down expression of this stress induced nuclear lncRNA. Using these ASOs we were able to successfully knock down LSINCT5 expression in the breast cancer cell line MCF7 and in the ovarian cancer cell line OVCAR5. The Figures below show the amount of knock-down obtained with the two LSINCT ASOs in MCF7 cells; also shown on this figure are the Cell Titre Blue Assay and BrDU staining showing a decrease in proliferation after LSINCT5 knock-down. This can be compared to the results using the NC ASO where there is good proliferation.



A: LSINCT5 knockdown utilizing LSINCT5-1 (L5-1) ASO (white bars) and LSINCT5-2 (L5-2) ASO (stripe bars) and NC ASO (black bars). **B:** Cell Titer Blue Assay showing decreased proliferation in L5-1 ASO (black bar) compared to NC ASO (white bar). **C:** BrdU staining (green) shows decrease proliferation in LSINCT5-1 ASO (bottom panel) compared to NC ASO (top panel) for 48 hours. Nucleus shown by DAPI stain (blue).

Attempts to knock down the expression of LSINCT5 further than 50% always resulted in cell death through apoptosis. As part of our future work we would like to try additional strategies (including adding apoptosis inhibitors) to see if we can decrease LSINCT5 expression below 50%.

In order to validate that the entire LSINCT5 transcript was being knocked down using the two ASOs synthesized, we created multiple qPCR primers in several regions of LSINCT5 and thus covering the entire non-redundant portion of the entire transcript. We did qPCR with these primers in MCF7 cells transfected with NC ASO and L5-1 ASO and showed that the entire region of LSINCT5 is indeed knocked down (see Figure 9A in the attached paper). We further used these primers with HMEC and MCF7 cells alone and saw that there is an increase in expression of the entire LSINCT5 region in the breast cancer cell lines as compared to HMEC.

We then examined the effects of decreased LSINCT5 expression on proliferation in both the breast and ovarian cancer cell lines and found a significant decrease of proliferation utilizing the LSINCT5-1 ASO by the cell titre blue assay. Even though we only detect a slight decrease of LSINCT5 expression in 24 hours with the LSINCT5-1 ASO, we still see a decrease in proliferation in the MCF7 cells as compared to normal controls. However, after 48 hours when there is considerably greater LSINCT5 knock-down, we see an even greater effect on cellular proliferation. In OVCAR5 cancer cells we notice no decrease in proliferation at 24 hours when LSINCT5 is not knocked down; however, at 48 hours we do see a decrease of proliferation when LSINCT5 is sufficiently knocked down.

To further characterize the effects of LSINCT5 expression on proliferation, we also assessed 5-bromo-2-deoxyurine (BrdU) incorporation after treating cells for 48hours with the LSINCT5-1 ASO. Knock down of LSINCT5 expression further significantly inhibited proliferation as shown by only 28% BrdU incorporation in MCF7 cells treated with the LSINCT5-1 ASO as compared to the ASO-scrambled oligonucleotide. All of the figures that illustrate each of these points are contained within the publication on the characterization of LSINCT5 which was recently published in *RNA Biology* (Silva et al. *RNA Biology* 2011).

We have therefore succeeded in completing most of what we proposed for the second Specific Aim for this proposal. We've found that even a 50% decrease in expression of LSINCT5 results in a measurably significant

decrease of proliferation in both the breast cancer cell line MCF7 and the ovarian cancer cell line OVCAR5. As part of our future goals we're going to see if the addition of apoptosis inhibitors will enable us to get even further knock-down of LSINCT5.

We also have several things that were proposed that we still need to do. This includes adding our full length LSINCT5 expression construct to cells to see what effect increased expression of LSINCT5 has on a "normal" human mammary epithelial cell line (HMEC). We have started some of these preliminary experiments but the problem we've had is that when transfected into HMEC (or other) cells, we saw a dramatic increase in LSINCT5 expression which was considerably greater than anything observed in vivo. The second thing that we proposed and that we still need to do is to generate stable transfectants with either LSINCT5 knock-down or over-expression. This is something that we're going to work on in the coming year.

Specific Aim #3: To characterize genes, non-coding transcripts and pathways that are affected by the modulation of LSINCT5 expression.

We will utilize the power of Next Generation DNA sequencing to characterize the effect of modulating the expression of LSINCT5 on the entire transcriptome of either normal breast epithelial cell lines or breast cancer cell lines. Hence, we will be able to characterize which coding genes are affected by modulating LSINCT5 expression as well as non-coding transcripts that have altered expression as a result of changing LSINCT5 expression. This work will enable us to determine how LSINCT5 is involved in cellular proliferation and what genes and pathways LSINCT5 interacts with.

The P.I. of this proposal, David I Smith, in addition to running his own research program is also the Chairman of the Technology Assessment Group for the Center for Individualized Medicine at the Mayo Clinic. As part of this responsibility, he has been working to develop the necessary infrastructure for Next Generation sequencing at the Mayo Clinic. The DNA Sequencing Core of the Mayo Clinic Rochester has been purchasing various Next Generation sequencing platforms including previously having two Illumina GAIIx sequencers. Currently the sequencing Core has 5 HiSeq 2000's and they've just obtained the Pacific Biosciences Single Molecular Sequencer. While Dr. Smith does not run this facility, he is still in charge of examining new technologies and advising the Mayo Clinic as to which platforms to examine and potentially purchase. Dr. Smith also has been conducting experiments with the Next Generation sequencers to evaluate their ability to address both research and its translation into clinical practice.

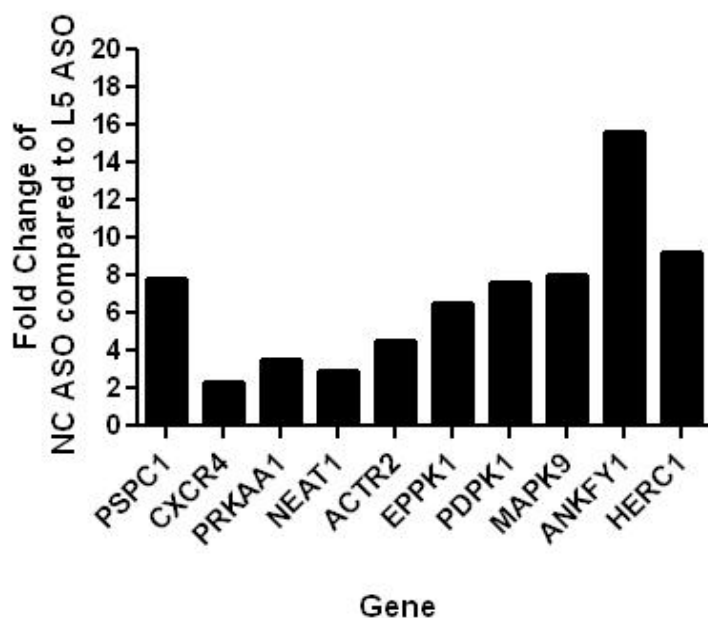
When this proposal was written, we had originally planned to utilize Next Generation sequencing to characterize any and all transcripts (including both coding and non-coding transcripts) whose expression was significantly changed when we decreased LSINCT5's expression. However, instead of using this very powerful new technology, we opted instead for testing for transcripts that had changed expression by hybridization to the Affymetrix U133 Plus2 Microarrays instead. The main reason for doing this was convenience and cost. The one disadvantage of doing this experiment this way was that it could only analyze some 20,000 genes, the vast majority of which are protein-coding genes. However, it should be mentioned that a number of important long non-coding transcripts, including NEAT1 and NEAT2, are actually on the Affymetrix U133 Plus 2 arrays.

When we compared the expression of transcripts when MCF7 cells were treated with the scrambled ASO nucleotide to those with a 50% decrease in LSINCT5 expression using the LSINCT5-1 ASO we found that there were 816 genes with either an increased or decreased expression of greater than 1.5 fold. Of these there were 96 genes whose expression changed at least two-fold (either up or down) in response to the LSINCT5 knock-down. The full list of these 96 genes are listed in our published paper (Silva et al. LSINCT5 is over expressed in breast and ovarian cancer and affects cellular proliferation *RNA Biology* 2011: 8(3).

We decided to select a subset of the 96 genes whose expression changed at least two-fold after knock-down of LSINCT5 for validation. Since LSINCT5 is localized in the nucleus, is over-expressed in several different

cancer types and whose over-expression appears to also increase proliferation, we attempted to validate the expression of genes that are associated with these criteria. We constructed qPCR primers for 36 of the 96 genes and were able to validate that 10 of these 36 genes did have more than a two-fold change in mRNA expression in response to the LSINCT5 knock-down.

The genes altered by LSINCT5 knockdown include a kinase (PDPK1), important nuclear assembly genes (NEAT1 and PSPC1), a gene involved in membrane transport (HERC1), a transcription factor (ANKFY1), a gene associated with carcinogenesis (EPPK1), stress (PRKAA1/AMPK), motility (ACTR1), T-cell differentiation, or a combination of these functions (CXCR4 and MAPK9/NKK2). The Figure below shows the changes in expression of these 10 genes in response to the 50% knockdown of LSINCT5 in the MCF7 breast cancer cell line.



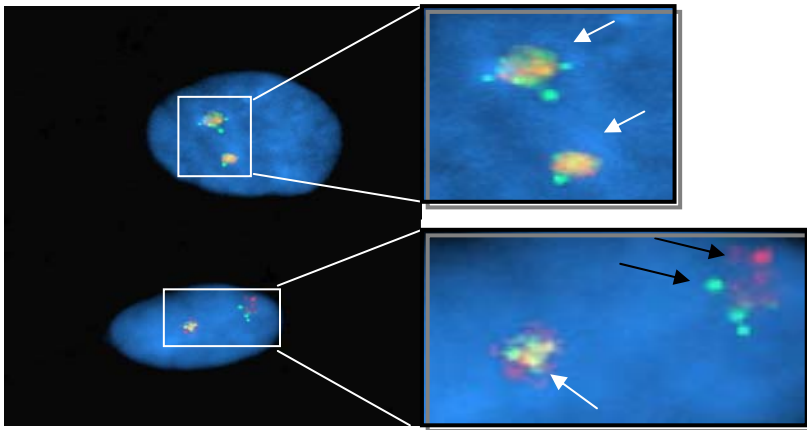
Two very interesting and important genes that had significant decreases in expression were the lncRNA NEAT1 and the protein-coding gene PSPC1 (paraspeckle component 1). It has previously been shown that NEAT1 is essential for the formation and maintenance of paraspeckles (Sunwoo et al. *Genome Research* 2009; 19: 347-359; Sasaki H. *Genome Biology* 2009; 10:227; Sakaki I et al., *PNAS* 2009; 106: 2525-2530) and that there are other both coding and non-coding transcripts that may be players in the formation of paraspeckles in addition to NEAT1 (Fox et al. *Current Biology* 2002; 12:13-25).

Another gene that had decreased expression after LSINCT5 knock-down was CXCR4, which is the most commonly used marker of several cancers and has also been identified as an important breast cancer marker associated with invasion and metastasis (Brand et al. *Exp Cell Res* 2005; 310: 117-130; Lee, *Oncogene* 2010; 29: 56-67; and Hassan S. *Int J Cancer* 2010; 129(1): 225-232). Our previous results have shown LSINCT5 to be over-expressed in metastatic breast cancer cell lines. These results suggest that since a decrease in LSINCT5 expression results in decreased CXCR4 expression, then quite possibly increased LSINCT5 expression (which is so commonly observed in breast and ovarian cancers) leads to increased expression of CXCR4, which may explain the proliferative effect when LSINCT5 expression is increased in breast and ovarian cancers.

We have therefore completed almost all of what we proposed in our original Breast Cancer Idea Proposal. However, there are additional questions that need to be addressed including the most important question of all, which is what is function of LSINCT5 in both normal cells and how does its over-expression affect the growth of developing breast cancer cells. These are important questions which need to be answered so that we will be

more competitive for extramural funding when we submit our proposal on the characterization of LSINCT5 and its role in the development of breast cancer to the National Cancer Institute.

To begin to address this important question we have started work in two areas that were not part of our original proposal, but which are extremely important questions that we need to address. The first was a set of experiments to examine the localization of LSINCT5 within the nucleus. Since LSINCT5 seemed to be linked to NEAT1 and paraspeckles, which are already known to interact with each other, we decided to start some RNA in situ hybridization experiments to examine LSINCT5 within cells. We therefore produced fluorescently labeled probes for RNA in situ hybridization for both LSINCT5 and NEAT1. When we optimized conditions for detecting LSINCT5 mRNA (which was a challenge because LSINCT5 had relatively low expression levels) we first found that the LSINCT5 was indeed localized in the nucleus, as expected from our nuclear and cytoplasmic extracts experiment. We then did RNA in situ hybridization with LSINCT5 and NEAT1 together. The figure below shows that these two do appear to co-localize, but these are very preliminary studies.



B. RNA in situ hybridization with MCF7 cells showing co-localization of LSINCT5 (green) and NEAT1 (red).

The second set of experiments (which we have just started) is to try and determine what protein molecules are associated with the LSINCT5 transcript. Our strategy to do this is to produce a biotin labeled LSINCT5 probe and then to hybridize this to nuclear extracts from MCF7 cells. Any proteins that are pulled down using streptavidin coated magnetic beads will be analyzed by the Mayo Clinic Proteomics Facility using one of their Mass Spectrometers to determine important proteins that might be involved with or interact with the LSINCT5 transcript.

KEY RESEARCH ACCOMPLISHMENTS

PUBLICATIONS AS A DIRECT RESULT OF THIS GRANT

- 1) Silva JM, Perez DS, Pritchett JR, Halling ML, Tang H, **Smith DI**. Identification of long stress-induced non-coding transcripts that have altered expression in cancer. *Genomics* 2010, 95(6): 355-362.
- 2) Silva JM, Boczek NJ, Berres MW, Ma X, **Smith DI**. LSINCT5 is over-expressed in breast and ovarian cancer and affects cellular proliferation. *RNA Biol* 2011: 8(3) [Epub ahead of print].

ABSTRACTS AS A DIRECT RESULT OF THIS GRANT

- 1) **Smith DI**, Bozcek N, Perez DS, and Silva J. LSINCT5, a long stress-responsive non-coding transcript is over-expressed in breast and ovarian cancers and promotes cellular proliferation. Submitted to and presented at the 2011 Era of Hope Breast Cancer Meeting in Orlando, Florida.

PRESENTATIONS RESULTING FROM THIS GRANT

I recently co-chaired a session at the 2011 Era of Hope Breast Cancer Meeting in Orlando on Non-coding RNA and Epigenetics in Breast Cancer. In addition, I gave a presentation on LSINCT5 (with the same title as the abstract listed right above).

LSINCT5 is overexpressed in breast and ovarian cancer and affects cellular proliferation

Jessica M. Silva,^{1,2} Nicole J. Boczek,¹ Michael W. Berres,¹ Xinghong Ma¹ and David I. Smith^{1,*}

¹Division of Experimental Pathology; Department of Laboratory Medicine and Pathology; ²Biochemistry and Molecular Biology; Mayo Clinic and Foundation; Rochester, MN USA

Key words: long ncRNA, cancer, proliferation, LSINCT5

Abbreviations: LSINCT, long stress induced non-coding RNA; lncRNA, long non-coding RNA; NEAT1, nuclear paraspeckle assembly transcript 1 (tncRNA); NEAT2, nuclear paraspeckle assembly transcript 1 (MALAT-1); PSPC1, paraspeckle component 1; XIST, X-inactive specific transcript; HOTAIR, hox transcript antisense RNA; PRINS, psoriasis susceptibility-related RNA gene induced by stress

More than 98% of the human genome is comprised of non-protein coding sequences. Interestingly, a considerable fraction of these sequences is transcribed into non-protein coding RNA transcripts. These transcripts range in size from very small RNAs such as the miRNAs (20–25 base pairs) to transcripts that can range up to 100 kb or more. Some longer non-coding RNAs (lncRNAs) have been found to play important regulatory roles within cells. In this report, we demonstrate that LSINCT5 is a 2.6 Kb polyadenylated, long stress-induced non-coding transcript that is on the negative strand, localized in the nucleus and potentially transcribed by RNA polymerase III. LSINCT5 is overexpressed in breast and ovarian cancer cell lines and tumor tissues, relative to their normal counterpart. In addition, knocking down the expression of LSINCT5 in cancer-derived cell lines causes a decrease in cellular proliferation. Finally, we identified 95 genes with more than 2-fold changes when knocking down LSINCT5 expression by using the Affymetrix U133 Plus 2 array. We chose a subset of these genes to validate using qPCR and found that ten of these genes were indeed significantly affected by the LSINCT5 knockdown. Interestingly, two genes that were significantly downregulated were the lncRNA NEAT-1 and a protein coding gene PSPC1. We have therefore characterized a novel lncRNA that is overexpressed in breast and ovarian cancers, enhances cellular proliferation and may play a significant role in multiple processes.

Introduction

The sequencing of the human genome revealed there are approximately 26,588 protein-coding genes.¹ However, the coding portion of the genome represents less than 2% of the genome. The remaining 98% of the genome consists of non-protein coding DNA sequences, including 45% middle- to highly-repetitive sequences and 53% non-repetitive sequences including introns, intergenic regions and regulatory RNA genes.² Whole genome tiling arrays are a powerful tool to examine transcription across the non-repetitive portion of the genome. These arrays contain oligonucleotide probes tiled across both the protein coding and the non-repetitive non-protein coding portions of the genome. Experiments with such arrays revealed that much of the non-repetitive portion of the human genome is transcriptionally active and that non-protein coding transcripts greatly exceed the number of protein coding transcripts.³

We now know that there are entire families of non-coding RNA transcripts. In particular, the miRNAs have been characterized as small ncRNAs that can modulate gene expression. Some of these transcripts are aberrantly regulated in diseases

such as HIV and cancer.^{4–6} Larger regulatory ncRNAs, referred to as long non-coding RNAs (lncRNAs), are longer than 200 nucleotides and have been shown to have diverse functions.⁷ For instance, lncRNAs can regulate expression of a downstream promoter in cis as is the case for *XIST* (X-inactive specific transcript) and X-inactivation,^{8,9} or a sequence on an entirely different chromosome, as is the case for *HOTAIR* (Hox transcript antisense RNA) transcript.^{8–10}

Some lncRNAs have been found to be aberrantly expressed in various diseases ranging from psoriasis to cancer.^{11,12} Two well-characterized, adjacent intergenic lncRNAs are *NEAT1* (tncRNA) and *NEAT2* (MALAT-1). *NEAT1* has been found to function as an essential structural determinant of paraspeckles.¹³ Currently, the formation of paraspeckles includes interactions between newly transcribed *NEAT1* RNA and DBHS (*Drosophila melanogaster* behavior, human splicing) protein dimers consisting of *PSPC1*, *P54NRB/NANO* or *PSF/SFPQ* in addition to RNA A to I hyperedited inverted repeats and potentially other RNAs and lncRNAs.^{14,15} *NEAT1* and *NEAT2* overexpression has been associated with several cancers including human hepatocellular carcinomas and hepatoblastomas, endometrial stromal

*Correspondence to: David I. Smith; Email: smith.david@mayo.edu

Submitted: 10/19/10; Revised: 12/13/10; Accepted: 12/15/10

DOI:

sarcoma of the uterus, cervical cancer, neuroblastoma and lung adenocarcinomas.^{16–19}

We previously utilized whole genome tiling arrays containing tiled oligonucleotides spaced 35 bp apart (center-center) to detect non-coding transcripts from normal human bronchial epithelial (NHBE) cells before or after treatment with the tobacco carcinogen nitrosamine 4-(methyl-nitrosamino)-1-(3-pyridyl)-1-butanol (NNK). We identified over 200,000 NNK stress-responsive transcripts which were further examined to focus on the identification of a small group of lncRNA stress-responsive transcripts (transcripts longer than 300 nucleotides that did not have open reading frames and had no homology to mRNAs or other protein coding transcripts). This work led to the identification of 12 such transcripts, which we termed *LSINCT5* (long stress induced non-coding transcripts). We previously described our preliminary characterization of these transcripts, noting that some were more abundantly expressed in proliferative normal human tissues.²⁰ In addition, we found that these 12 lncRNAs were overexpressed in some lung and breast cancers relative to normal epithelial cells.

In this report, we describe our analysis of one of the lncRNAs, *LSINCT5*. *LSINCT5* is an intergenic lncRNA (chr5: 2,765,705–2,768,351) that is located 829,825 bp downstream of the *IRX4* gene and 31,529 bp upstream of the *IRX2* gene (information taken from UCSC Browser March 2006 version). *LSINCT5* does not have an open reading frame or any homology to any mRNA or genes as previously shown.²⁰ This transcript was specifically chosen because it was greatly overexpressed in most of the breast cancer cell lines examined. We now demonstrate that *LSINCT5* is a polyadenylated 2,674 bp transcript that is transcribed from the negative strand potentially by RNA Polymerase III. We also found that *LSINCT5* had increased expression in breast and ovarian cancer cell line and tumor tissue panels and demonstrate that *LSINCT5* knockdown impairs cellular proliferation. Finally, we identified a number of genes whose expression significantly changes in response to knockdown of *LSINCT5* expression, several of which could play important roles in cancer development.

Results

Characterization of *LSINCT5*. *LSINCT5* transcript length and direction. In a previous report, we demonstrated using northern blotting that *LSINCT5* is produced as an approximately 3 Kb transcript.²⁰ The tiling array experiment identified a 344 bp core sequence defined by adjacent oligonucleotide probes which hybridized more strongly to RNA derived from NHBE cells treated with NNK (Fig. 1B termed L5 tiling array). In order to identify the precise size as well as where this transcript begins and ends, we conducted 5' and 3'RACE using primers derived every 500 bp both upstream and downstream from the L5 tiling array 344 bp core sequence. Utilizing one of the upstream primers (L5 5'RACE primer, sequence listed in Sup. Table 2 and location shown in Fig. 1B) located at the beginning of the core sequence at chr5: 2,768,351–2,768,371, we identified a 130 bps product for 5'RACE (Fig. 1A). We also amplified a 500 bps product with one of the downstream primers (L5 3'RACE primer, sequence listed in Sup. Table 2 and location shown in Fig. 1B) located at

chr5: 2,765,705–2,765,761, 2,323 bp of the core sequence for 3' RACE of the *LSINCT5* transcript (Fig. 1A). The 5'RACE and 3' RACE products were isolated and verified by sequencing. All together this defines a transcript of precisely 2,647 nucleotides on chr5: 2,765,705–2,768,351 shown in Figure 1B (sequence listed as GenBank locus GU228577 and in Sup. Fig. 1).

We observed more *LSINCT5* expression when utilizing polyA+ to prime cDNA synthesis as compared to random primed cDNA produced from total RNA (Sup. Fig. 2). We also identified a polyadenylation signal downstream of the identified 3' end of the full length transcript (Sup. Fig. 1). Further experimentation will be needed to provide clear evidence of the transcripts polyA tail length. The results from 5' RACE and 3' RACE and the enrichment of *LSINCT5* transcripts in poly A+-selected cDNA suggests that *LSINCT5* is a polyadenylated lncRNA.

In order to identify and confirm the directionality of this transcript, we cloned the *LSINCT5* transcript sequence into a directional dual promoter (SP6/T7 polymerase) plasmid and expressed either promoter for the positive strand or negative strand RNA transcription for hybridization onto northern blots. We made northern blots containing total RNA from a breast cancer cell line panel and hybridized the blots first with a positive strand-derived probe and then a negative strand probe derived from the dual promoter plasmid. We only detect hybridization with the probe complementary to the negative strand (Fig. 1C). The northern blots also demonstrate considerably higher expression of *LSINCT5* in the breast cancer cell lines as compared to the normal breast epithelial cell line. We also show *LSINCT5* to now contain two LINE elements as shown in Figure 1B and Supplemental Figure 1, which are also in the negative strand direction (indicated by arrows).

***LSINCT5* localization.** In order to identify the intracellular location of *LSINCT5*, we isolated total RNA, nuclear RNA and cytoplasmic RNA lysates from a breast cancer cell line, converted the RNA into cDNA and then used quantitative real-time PCR (qPCR) to measure *LSINCT5* transcripts in the fractionated RNAs. As a positive control we quantified the expression of three genes, *GAPDH* which is found in both the nucleus and cytoplasm, *U6* a nuclear transcript involved in RNA processing and a mitochondrial gene *MT RNRI* for use as a cytoplasmic control. *LSINCT5* demonstrated significantly increased expression in the nuclear fraction, while the three control genes listed above demonstrated the expected expression patterns (Fig. 2A). These results suggest *LSINCT5* to have higher expression in the nucleus.

***LSINCT5* transcription.** Because *LSINCT5* is a novel lncRNA, we characterized which RNA polymerase might be responsible for its transcription. We treated BT474 cells with a RNA polymerase II inhibitor, α -amanitin or an RNA polymerase III inhibitor, RNAP III inhibitor. We then used qPCR to quantify expression levels of *LSINCT5* relative to two control transcripts, *5S rRNA* (transcribed by RNA polymerase III) and *cMYC* (transcribed by RNA polymerase II). Results are shown in Figure 2B. As expected, *5S rRNA* expression was decreased in cells treated with RNAP III inhibitor and unchanged in cells treated with α -amanitin. Conversely, *cMYC* expression was decreased in cells

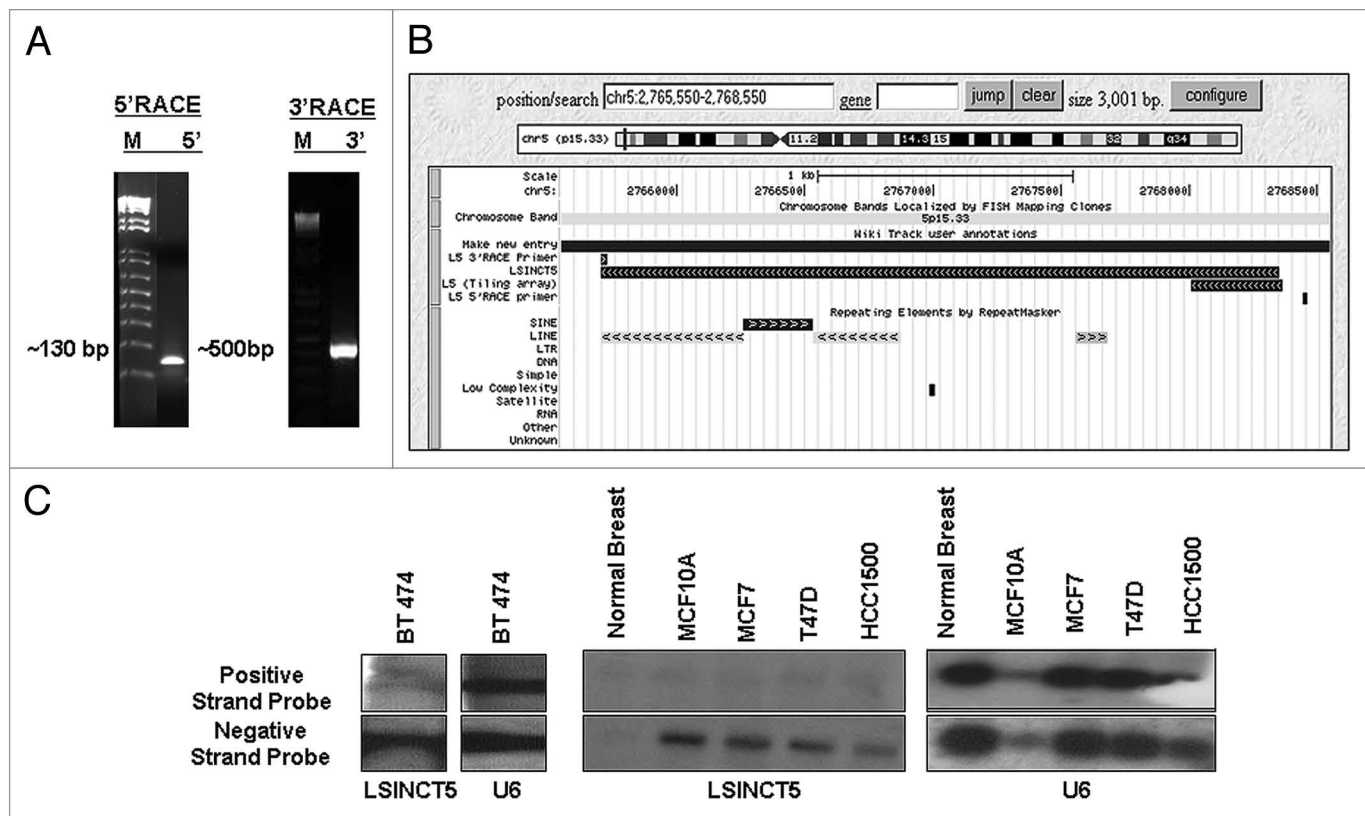


Figure 1. LSINCT5 is a long ncRNA that is transcribed on the negative strand. (A) LSINCT5 RACE products showing 5'RACE and 3' RACE PCR bands. (B) LSINCT5 transcript shown by the UCSC Genome Bioinformatics Browser. UCSC Genome Browser shows LSINCT5 on Chromosome band 5p15.33. Wiki track shows the location of the LSINCT5 3'RACE primer (L5 3'RACE) and 5'RACE primer (L5 5'RACE), LSINCT5 full length (LSINCT5) and LSINCT5 region identified from previous tiling array (L5 tiling array). Repetitive elements by RepeatMaster are also shown indicating direction by arrows. LSINCT5 is a 2 647 bps lncRNA that contains two LINE elements. (C) Northern Blot showing LSINCT5 transcript utilizing directional dual promoter plasmid in a breast cancer cell line RNA part and one normal breast RNA. U6 was utilized as a positive housekeeping gene. LSINCT5 was expressed only on the negative direction in the breast cancer RNAs.

treated with α -amanitin and unchanged in cells treated with RNAP III inhibitor. *LSINCT5* expression was decreased only in cells treated with RNAP III inhibitor and was unchanged in cells treated with α -amanitin, suggesting that *LSINCT5* is potentially transcribed by RNA Polymerase III.

***LSINCT5* has increased expression in breast and ovarian cancer cell lines.** We previously reported low *LSINCT5* expression in less proliferative normal tissues such as brain and higher levels of expression of *LSINCT5* in highly proliferative normal tissues including kidney, pancreas, colon and spleen (Sup. Fig. 3). Previously we demonstrated that *LSINCT5* had increased expression in a small panel of breast cancer and lung cancer cell lines relative to normal epithelial-derived cell lines. In this report, we examined the expression of *LSINCT5* in a larger number of breast cancer cell lines and also in several ovarian cancer cell lines (Fig. 3A and B). *LSINCT5* has greater than ten fold increased expression in all cancer cell lines tested as compared to normal cell lines from the same tissues. Several breast cancer cell lines had 30-fold higher expression of *LSINCT5* than in HMEC, including MDA468, T47D and BT474 cells (Fig. 3A). *LSINCT5* was also overexpressed by more than 10 fold in the ovarian cancer cell lines tested as compared to normal ovarian epithelium (Fig. 3B).

***LSINCT5* expression in breast and ovarian primary tumor tissues.** We then examined the expression of *LSINCT5* in fresh frozen tumor tissue samples derived from breast and ovarian cancers and found increased *LSINCT5* expression in both tumors tissue panels (Fig. 3C and D). *LSINCT5* was greatly overexpressed between two to seven fold in the primary breast cancers as compared to normal benign breast tissue (Fig. 3C). *LSINCT5* also demonstrated between 5 to 26 fold increased expression in primary ovarian tumors as compared to normal ovarian tissue (Fig. 3D).

Knocking down *LSINCT5* expression decreases proliferation in breast and ovarian cancer cell lines. *LSINCT5* knock-down in cancer cells. The analysis of *LSINCT5* expression in the various cancers and tumor tissue panels indicates that *LSINCT5* could play a role in carcinogenesis. In order to identify *LSINCT5*'s potential function, we examined whether *LSINCT5* promotes cellular proliferation. The use of antisense oligos (ASOs) has recently been shown to successfully knock down the expression of nuclear ncRNAs.²¹ We created two ASOs specific to *LSINCT5* (LSINCT5-1 and LSINCT5-2 ASOs) and a scrambled Normal Control (NC ASO) to knock down expression of this stress induced nuclear lncRNA (sequences listed in Sup. Table 2). We

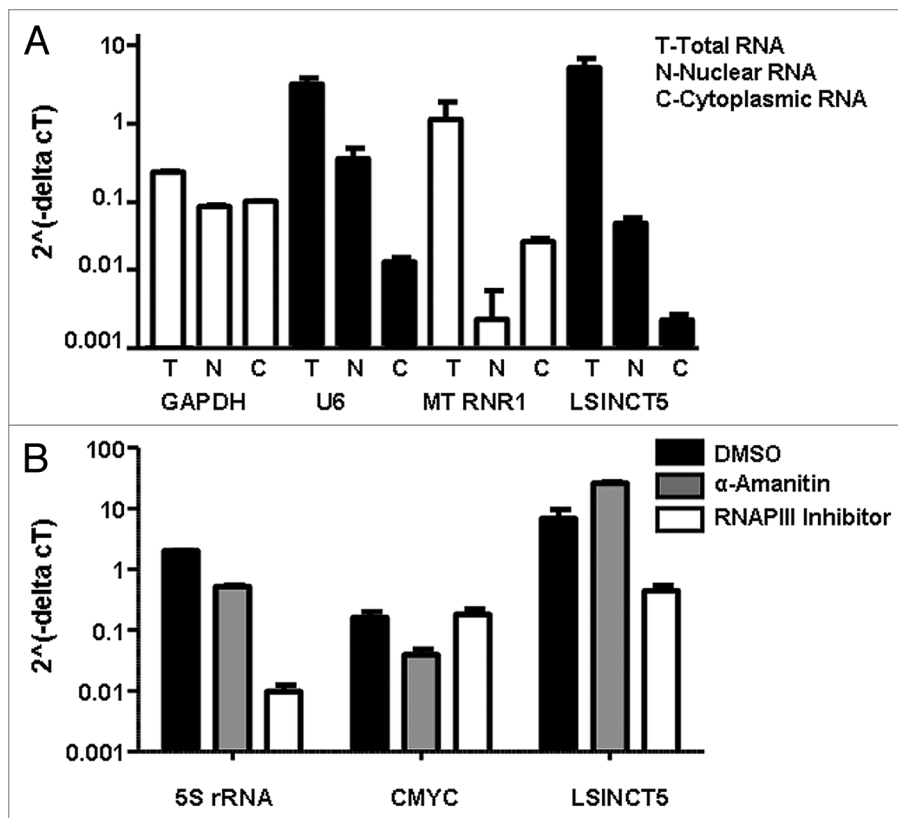


Figure 2. LSINCT5 is localized in the nucleus and transcribed by RNA polymerase III. QPCR analysis on (A) Three control genes and LSINCT5 expression in total (T) RNA, nuclear (N) RNA or cytoplasmic (C) RNA. Housekeeping gene Actin was used for normalization. LSINCT5 is expressed more in the nuclear RNA. (B) Control genes and LSINCT5 expression when treating with α -amanitin (RNA Polymerase II inhibitor) or RNAP III inhibitor (RNA Polymerase III inhibitor) in BT474 cells, utilizing U6 and Actin housekeeping genes for normalization, respectively. QPCR experiments were conducted in triplicate and repeated two times.

knocked down *LSINCT5* expression in both the breast cancer cell line MCF7 (Fig. 4A) and the ovarian cancer cell line OVCAR5 (Sup. Fig. 4A). In MCF7 cancer cells, we tried different concentrations and times of knockdown with the two distinct ASOs and found the greatest observable knock-down of *LSINCT5* expression in 24 hours with *LSINCT5*-1 ASO and in forty-eight hours with the *LSINCT5*-2 ASO at a concentration of 50 pM. In OVCAR5 cancer cell lines, we observe *LSINCT5* knockdown with only the *LSINCT5*-1 ASO in forty-eight hours (Sup. Fig. 4A). Our optimal conditions for *LSINCT5* knock-down produced only 55% and 50% knockdown of *LSINCT5* expression in MCF7 and OVCAR5 cancer cells respectively. Attempts to get greater knock-down resulted in apoptosis in the cancer cell lines.

***LSINCT5* knockdown decreases cellular proliferation.** We examined the effects of decreased *LSINCT5* expression on proliferation in both the breast and ovarian cancer cell lines and found a significant decrease of proliferation utilizing *LSINCT5*-1 ASO by cell titer blue assay (Fig. 4B and Sup. Fig. 4B). Although in twenty fours we only slightly decrease *LSINCT5* expression with the *LSINCT5*-1 ASO we still see a decrease of proliferation in the MCF7 cells compared to normal controls. However, after 48 hours, when there is considerably greater *LSINCT5* knock-down

we see an even greater effect on cellular proliferation (Fig. 4B).

To further characterize effects of *LSINCT5* on proliferation, we also assessed 5-bromo-2-deoxyuridine (BrdU) incorporation after treating cells in forty-eight hours with the *LSINCT5*-1 ASO and NC ASO. Knock down of *LSINCT5* expression further inhibited proliferation as shown by only 28% BrdU incorporation in cells treated with the *LSINCT5*-1 ASO as compared to the ASO-scrambled oligonucleotide (Fig. 4C and Sup. Fig. 5).

Genes identified when knocking down *LSINCT5* expression. *LSINCT5* is a cancer-associated lncRNA which plays a role in proliferation. In order to further characterize *LSINCT5*'s function within cells, we hybridized RNA isolated from the *LSINCT5* knockdown from MCF7 cells to Affymetrix gene expression microarrays as compared to RNA isolated from the scrambled normal control oligo. There were 816 genes with either increased or decreased expression greater than 1.5 fold in response to the 55% decrease of *LSINCT5* expression in MCF7 cells. There were 96 genes whose expression changed at least two-fold (Sup. Table 1). Because *LSINCT5* is localized in the nucleus and is overexpressed in cancer and also increases proliferation, we attempted to validate genes associated with these criteria. As shown in Table 1, we performed qPCR on 36 selected genes (primer

sequences listed in Sup. Table 2) and were able to validate that 10 of these genes (Fig. 5) had more than two-fold increase or decrease of expression in response to the *LSINCT5* knockdown.

The genes altered by *LSINCT5* knockdown include several kinases (PDPK1), important nuclear assembly genes (NEAT1 and PSPC1), membrane transport process genes (HERC1), a transcription factor (ANKF41) and genes associated with carcinogenesis (EPPK1), stress (PRKAA1/AMPK), motility (ACTR2), T-cell differentiation or a combination of these functions (CXCR4 and MAPK9/JNK2). The significant gene expression changes due to the decrease of *LSINCT5* expression demonstrates *LSINCT5*'s important role in many cellular processes including proliferation.

Discussion

The discovery of multiple functional regulatory ncRNAs has lead to genome-wide searches in multiple species as well as for transcripts that are aberrantly expressed in various disease states. While most of these studies have focused on the identification of smaller ncRNAs such as miRNAs, there are many lncRNAs which have yet to be identified and characterized.

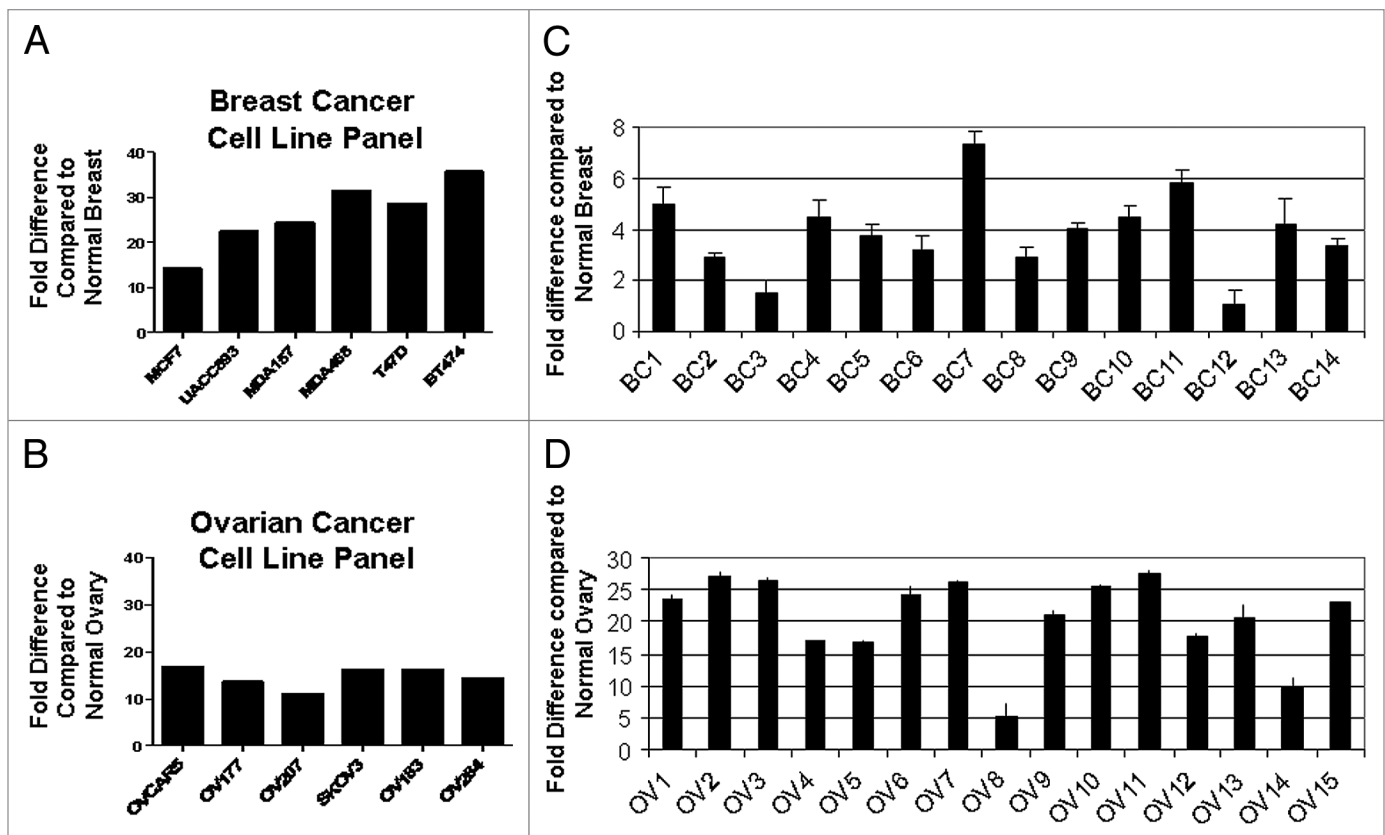


Figure 3. LSINCT5 expression is increased in breast and ovarian cancer cell lines and tumor tissues. LSINCT5 fold difference in (A) breast cancer cell line part, (B) ovarian cancer cell line part (C) breast primary tumor tissue part (tissues labeled BC) and (D) ovarian primary tumor tissue part (tissues labeled OV) compared to normal counterparts. QPCR experiments were conducted in triplicate and repeated two times.

In this paper we successfully characterize one lncRNA, *LSINCT5*, a long stress-induced non-coding transcript. The *LSINCT5* transcript consists of a core sequence and two long interspersed elements (LINE 1M repeat and LINE 1ME repeat). These repeat elements are not uncommon in ncRNAs and are found in the well-known *XIST* ncRNA. The LINE elements in *XIST* have been shown to facilitate the assembly of a heterochromatic nuclear compartment induced by *Xist* and recruit genes for silencing the X chromosome inactivation center region.²² The significance of common DNA repeat elements within the *LSINCT5* transcript is currently unknown however results from our gene analysis when knocking down *LSINCT5* suggest it too may function in recruiting genes.

In addition to characterizing the *LSINCT5* transcript, we demonstrate that this RNA is polyadenylated and potentially transcribed by RNA Polymerase III (RNAP III). We show *LSINCT5* to be more highly expressed in oligodT primed cDNA as compared to random primed cDNA and also shown that it contains a polyadenylation signal by its 3' end. Several RNAP III transcribed ncRNAs are well characterized; these include *tRNA* (decodes mRNA sequence into the order of amino acid residues in proteins), *5S rRNA* (essential component of ribosome's), *7SL RNA* (trafficking of nascent polypeptides to membranes) and *BC200 RNA* (human counterpart of *BC1 RNA* a neuronal specific RNA). These RNAP III transcribed ncRNAs have also been

shown to have altered expression in cancers including breast and ovarian carcinomas similar to *LSINCT5*.²³⁻²⁵ These results are preliminary results and further critical experimentation will need to be conducted to finalize these interesting findings, especially that this novel long non-coding transcript is indeed transcribed by RNA polymerase III. If true, this would indeed be novel as most RNA polymerase III transcripts are considerably shorter.

The list of lncRNAs associated with stress and cancer is increasing. For example, an interesting stress responsive lncRNA which may also be associated with cancer is *PRINS* (psoriasis susceptibility-related RNA gene induced by stress). *PRINS* was found to play a protective role in cells exposed to stress, furthermore, elevated *PRINS* expression in the epidermis may contribute to psoriasis susceptibility.²⁶ *PRINS* has also been found to regulate the overexpressed anti-apoptotic protein G1P3, an interferon-inducible gene with anti-apoptotic effects in cancer cells.¹¹ We utilized a whole genome tiling array for analysis of RNA from NHBE cells treated with NNK to identify the stress-responsive cancer associated lncRNA, *LSINCT5*. The tobacco carcinogen NNK causes DNA mutations that dysregulate various genes affecting viability, cell movement, cell cycle and cell proliferation in cancer cells.²⁷⁻³¹ *LSINCT5* is a new addition to the stress and cancer associated lncRNA family. *LSINCT5* has increased expression in multiple cancers and tumor tissues including breast and ovarian, compared to their normal counterparts. We also

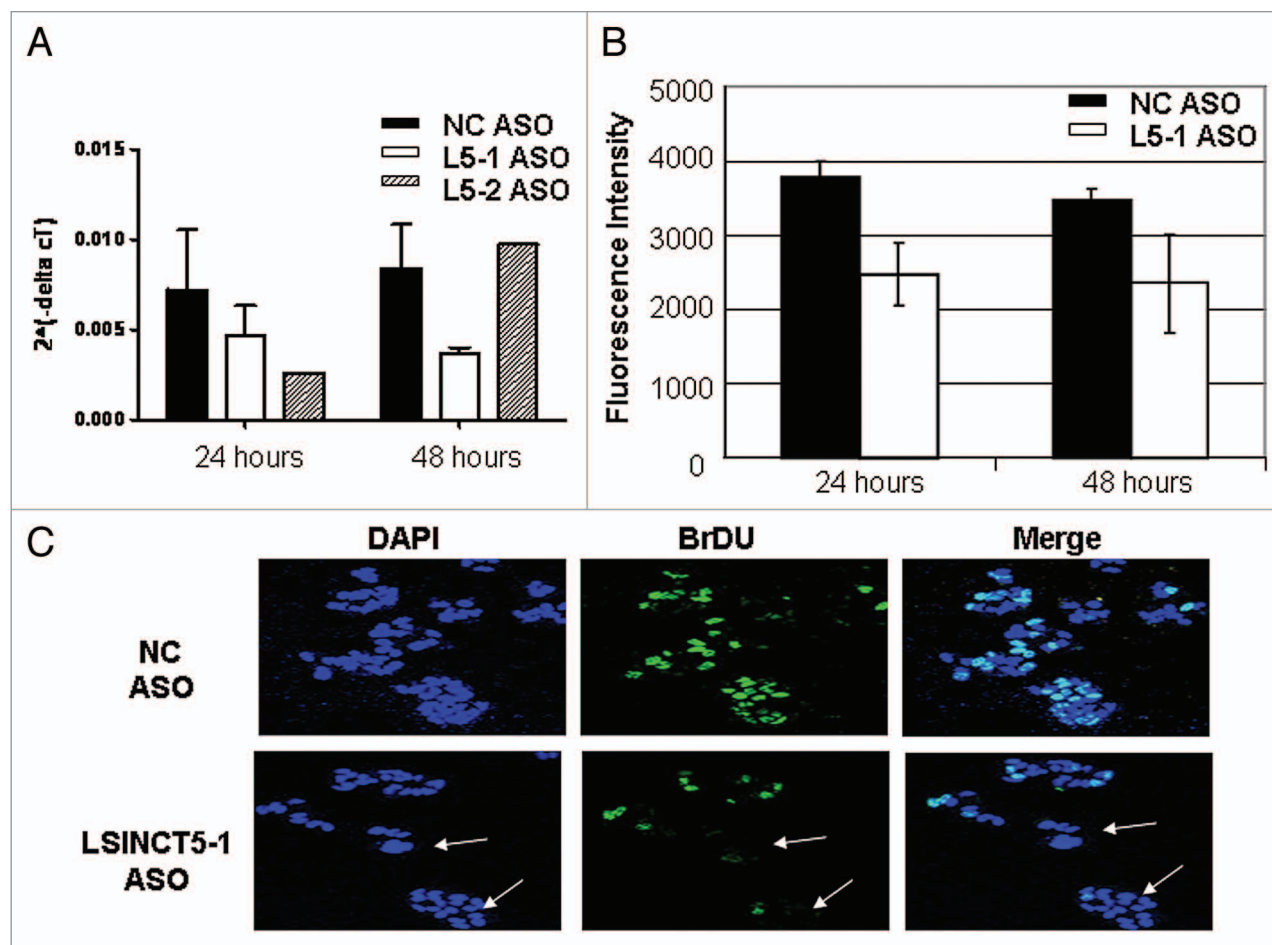


Figure 4. LSINCT5 knockdown causes a decrease in proliferation in MCF7 cancer cells. (A) Knockdown of LSINCT5 with NC ASO (black bars), LSINCT5-1 (L5-1) ASO (white bars) and LSINCT5-2 (L5-2) ASO (striped bars) for twenty-four and forty eight hours. (B) MCF7 cancer cells showing a decrease of proliferation in LSINCT5-1 ASO (white bars) when compared to NC ASO (black bars) in twenty-four and forty eight hours. (C) NC ASO (top part) and LSINCT5-1 ASO (bottom part) showing a decreased of proliferation by fewer BrDU (green) staining in MCF7 cells in forty-eight hours. Nucleus is stained with DAPI (blue). BrDU experiment was conducted in duplicate and repeated two times.

find increased expression of *LSINCT5* in cervical cancer cell lines (data not shown). Although *LSINCT5* has increased expression in breast, ovarian and some cervical cancers it is not highly over-expressed when examining either lung cancer cell lines or primary endometrial tumors (data not shown) indicating some level of tissue specificity.

The increased expression of *LSINCT5* in normal tissues that are highly proliferative and increased expression in certain cancers suggests a possible role of *LSINCT5* in proliferation. When we knocked down expression of *LSINCT5*, this caused a decrease of proliferation in both a breast and ovarian cancer cell line. *NEAT1* and *NEAT2* are the two characterized lncRNAs that have also been found to play a role in proliferation. Our previous tiling array experiment with NHBE cells treated with NNK demonstrated (data not shown) that *NEAT1* and *NEAT2* are also stress responsive transcripts. Our results with *LSINCT5* would suggest that this is also an important lncRNA whose increased expression plays a role in cellular proliferation.

Finally, we identified a large number of genes that were altered when *LSINCT5* expression was decreased. We tested a sub-set

of these genes based upon important cancer and proliferation-related criteria and of 36 genes chosen we validated 10 that had greater than a two-fold change when *LSINCT5* expression was decreased 55%. Two very interesting and important genes that had significant decreases in expression were the lncRNA *NEAT1* and the protein-coding gene *PSPC1* (paraspeckle component 1). It's been previously shown that *NEAT1* is essential for the formation of and maintenance of paraspeckles³²⁻³⁴ and that there are other both coding and non-coding transcripts that may be players in the formation of paraspeckles in addition to *NEAT1*.¹⁵ We hypothesize *LSINCT5* to potentially be one of these players. These results suggest a potential role of *LSINCT5* in the paraspeckle formation. Currently, we are in the process of proving our hypothesis by utilizing RNA fluorescent in-situ hybridization. Another gene that had decreased expression after *LSINCT5* knock-down was *CXCR4*, which is the most commonly used marker of several cancers and has also been identified as an important breast cancer marker associated with invasion and metastasis.³⁵⁻³⁷ Our previous results have shown *LSINCT5* to be overexpressed in metastatic breast cancer cell lines. These results suggest that decreased

LSINCT5 leads to decreased CXCR4 or increased LSINCT5 leads to increased CXCR4, which may explain its proliferative effect. Experiments are currently being conducted to test these hypotheses.

LSINCT5, the stress-regulated lncRNA is a novel nuclear gene that appears to play a role in cellular proliferation and also with the development of breast and ovarian cancers.

Materials and Methods

Integrated Genome Browser (IGB). The data from the genome tiling array was analyzed using the Integrated Genome Browser (IGB) software to review expression of LSINCT5. The July v34 chip analysis was used set at a p value of < 0.01 with a minimum value of 300 base pairs.

University of California Santa Cruz genome bioinformatics browser (UCSC genome browser). UCSC *Homo sapiens* genome browser gateway March 2006 (NCBI Build 36.1) was used to analyze the LSINCT region was accessed at <http://genome.ucsc.edu>. We first utilized Genome Blat Search in order to validate isolated DNA sequences from RACE experiment. This analysis also included several of the UCSC genome browser fields. The Wiki track user annotations browser was used to submit LSINCT5 transcript sequence and primer sequences used for RACE experiments for validating genome location and for easy visualization of primer sequence in reference to LSINCT5. The Repeating Elements by RepeatMaster was also used to show repeats in the LSINCT5 region.

Cell lines, cell culture and tissues. All cell lines were grown in 5% CO₂ at 37°C incubator. HMEC cells were purchased from Lonza (CC-2551) and grown in MEBM medium (CC-3150). MCF7 cell line (HTB-22) was grown in Gibco Dulbecco's modified eagle medium (21063-029), media with Sigma L-glutamine (G7513) and Sigma pyruvate (S8636), 10% Sigma FBS (F4135) and supplemented with 0.1 mM CellGro MEM non-essential amino acids (25-025-CL), 1x Gibco Antibiotic/Antimycotic (15240), Sigma insulin solution (I9278) and 10 mM Sigma 17-estradiol β (E2257). MCF10A (CRL-10317) and MDA157 (HTB-24) were grown in Gibco DMEM with Sigma L-glutamine and 10% Sigma FBS. The MDA468 (HTB-132) cell line was cultured in Invitrogen L15 medium (11415-064) with Sigma L-glutamine and 10% Sigma FBS. T47D (HTB-133), SCOVI (HTB-77), HCC1500 (CRL-2329), BT474 (HTB-20) and UACC893 (CRL-1902) were purchased from ATCC. All cells were maintained according to providers' recommendations. OVCAR5 cell line was an established NIH cell line.³⁸ All other ovarian cancer cell lines were isolated and cultured as instructed.³⁹ Cryopreserved NHBE cells were purchased from Cambrex Bio Science (CC-2540) and grown in defined Cambrex BEGM Bullet kit medium with supplements (CC-3170) at manufacturer-suggested concentrations. Stress-induced cells NHBE cells were prepared by treating with 200 μM NNK (4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone) (M325750) from Toronto Research Chemicals prepared in dimethyl sulfoxide for 24 hours.

All tissues used in this study were collected from patients at the Mayo Clinic, Rochester, MN. All tissues were snap frozen in

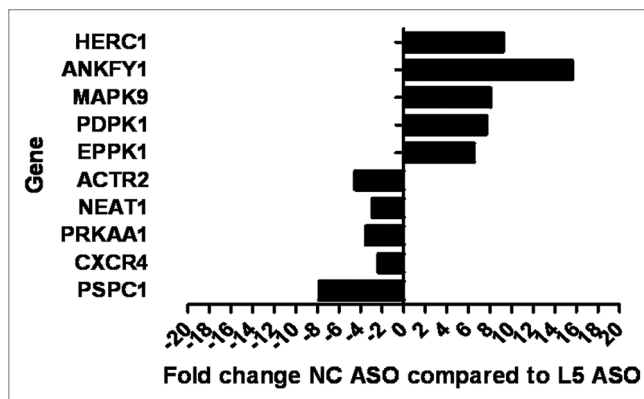


Figure 5. Identified genes regulated by LSINCT5. QPCR was conducted on 36 genes identified from a gene expression array on MCF7 cells transfected with NC ASO and L5-1 ASO at optimal concentration. 10 of the selected 36 genes were validated to have more than two-fold increase or decrease in expression when LSINCT5 was knocked down in MCF7 cells. The qPCR experiment was conducted in triplicate and repeated two times.

liquid nitrogen and stored according to the procedure approved by the Mayo Clinic Institutional Review Board. Breast and ovarian tumor tissues were compared to three benign breast tissues or normal ovaries for analyzing fold difference by qPCR.

Northern blots. Primers for *LSINCT5* were developed using Primer3 to be used for real time PCR experiments in order to create probes for subsequent northern Blot experiments. Probes were labeled with ³²P isotopes using GE Healthcare Megaprime DNA Labeling Systems kit (RPN1604). RNA (2–5 μg) was run on a 1% agarose gel containing 5% Sigma formaldehyde solution (F8775). RNA was transferred onto a nitrocellulose membrane. The membranes containing RNA were incubated with the labeled probes overnight using Ambion ULTRAhyb buffer (AM8670). Following day, membranes were washed and exposed for 1 to 3 days.

Directional northern blots. *LSINCT5* dual promoter plasmid was created by cloning real-time PCR amplified *LSINCT5* DNA, as stated in real-time PCR methods section, into the Invitrogen Dual Promoter Vector (45-0640) as instructed in TOPO TA cloning protocol. Several colonies were picked from the transformation and sequenced for validation of integrated *LSINCT5*. Ambion Maxiscript SP6/T7 kit (AM1322) was used to create negative and positive direction probes for *LSINCT5* to identify transcriptional direction by northern blot. Briefly, *LSINCT5* dual promoter plasmid is incubated with dATP, dUTP, dGTP and labeled [³²P]-dCTP and SP6 or T7 phage RNA polymerase enzyme at 37 degrees for 1 hour, TURBO DNASE is added and finally the reaction is stopped by adding EDTA. Probes are then used for northern Blots (as stated above in northern Blots section).

Ambion first choice RNA ligase mediated rapid amplification of cDNA ends (RLM RACE) (AM1700). 5' RACE. Total selected RNA is treated with Calf Intestine Alkaline Phosphatase (CIP) removing free 5'-phosphates; the RNA is then treated with Tobacco Acid Pyrophosphatase (TAP) to remove the cap structure

Table 1. Selected list of genes associated with proliferation and cancer, including nuclear genes altered by LSINCT5 knockdown identified by gene expression microarray

Representative public ID	Gene symbol	Gene title	Microarray fold change	qPCR fold change
AA969958*	PSPC1	Paraspeckle Component 1	-4.13	-7.82
L49506	CCNG2	Cyclin G ₂	-3.40	NV
BE675435	KLF6	Kruppel-Like Factor 6	-3.36	-1.55
BC020691	NAMPT	Nicotinamide Phosphoribosyltransferase	-2.46	-0.54
AF111116	BAG4	BCL2-Associated Athanogene 4	-2.46	NV
AF119855	PTPN11	Protein Tyrosine Phosphatase, Non-Receptor Type 11	-2.35	NV
Z25422	STK3	Serine/Threonine Kinase 3 (STE20 Homolog, Yeast)	-2.34	-0.57
AL553320	STIP1	Stress-Induced-Phosphoprotein 1	-2.26	NV
NM_004290	RNF14	Ring Finger Protein 14	-2.07	NV
BC017770	RBM8	RNA Binding Motif Protein 8	-2.10	NV
AL533352	RHBDD2	Rhomboid Domain Containing 2	-2.04	NV
NM_006572	GNA13	Guanine Nucleotide Binding Protein (G Protein), Alpha13	-2.03	-0.288
AF019888	ARPC4	Actin Related Protein 2/3 Complex, Subunit 4, 20 kda	-2.01	NV
AF100763*	PRKAA1	Protein Kinase, AMP-Activated, Alpha1 Catalytic Subunit	-2.01	-3.5
NM_001274	CHEK1	CHK1 Checkpoint Homolog (<i>S. Pombe</i>)	-2.01	NV
U32974	XIAP	X-Linked Inhibitor Of Apoptosis	-2.00	NV
AF132202	MALAT1	Metastasis Associated Lung Adenocarcinoma Transcript 1 (Non-Protein Coding)	-1.98	-1.93
J03263	LAMP1	Lysosomal-Associated Membrane Protein 1	-1.87	NV
L01639*	CXCR4	Chemokine (C-X-C Motif) Receptor 4	-1.86	-2.32
AU155361*	NEAT1	Nuclear Paraspeckle Assembly Transcript 1 (Non-Protein Coding)	-1.81	-2.89
BU175810*	ACTR2	ARP2 Actin-Related Protein 2 Homolog (Yeast)	-1.79	4.48
AL561296	E2F2	E2F Transcription Factor 2	-1.69	NV
NM_004162	RAB5A	RAB5A, Member RAS Oncogene Family	-1.68	NV
BC038960	HERC3	Hect Domain And RLD 3	-1.59	NV
AF061812	KRT16	Keratin 16	-1.59	-1.51
M57763	ARF6	ADP-Ribosylation Factor 6	-1.57	NV
AL569511	KRT6A	Keratin 6A	-1.54	-1.53
AL569511	KRT6C	Keratin 6C	-1.54	-1.53
BE622627	PIK3R3	Phosphoinositide-3-Kinase, Regulatory Subunit 3 (Gamma)	1.55	-0.175
AK025960*	ANKFY1	Ankyrin Repeat And FYVE Domain Containing 1	1.55	15.62
AI798790	DST	Dystonin	1.60	1.31
AL040394*	PDPK1	3-Phosphoinositide Dependent Protein Kinase-1	1.62	7.66
AL137725*	EPPK1	Epiplakin 1	1.62	6.52
AI808345*	MAPK9	Mitogen-Activated Protein Kinase 9	1.72	8.05
NM_003922*	HERC1	Hect (Homologous To The E6-AP (UBE3A) Carboxyl Terminus) Domain And RCC1 (CHC1)-Like Domain (RLD) 1	2.29	-9.29
NM_004667	HERC2	Hect Domain And RLD 2	2.30	NV
AB040875	SLC7A11	Solute Carrier Family 7, (Cationic Amino Acid Transporter, Y+ip System) Member 11	2.38	NV

*Significant genes with more than 2 fold change by qPCR. NV, Not Validated.

from full-length mRNA, leaving a 5'-monophosphate. A 45 base RNA Adapter oligonucleotide is ligated to the RNA population using T4 RNA ligase. A random-primed reverse transcription

reaction and nested PCR then amplifies the 5' end of a specific transcript. Ambion provides two nested primers corresponding to the 5' RACE Adapter sequence, and the user supplies two nested

antisense primers specific to the target gene. Sequence for LSINCT5 5'RACE used to identify LSINCT5 5'end is listed in **Supplemental Table 2** and location for primer is shown in **Figure 1B**.

3' RACE. The first strand cDNA is synthesized from total RNA, using the supplied 3' RACE adapter. The cDNA is then subjected to PCR using one of the 3' RACE Primers which are complimentary to the anchored adapter, and a user-supplied primer for the gene-of-interest. Sequence for LSINCT5 3'RACE used to identify LSINCT5 3'end is listed in **Supplemental Table 2** and location for primer is shown in **Figure 1B**.

Bands identified from both 5' and 3' RACE real-time PCR were sequenced for validation.

Real-time polymerase chain reaction (real-time PCR). Real-time PCR was used to analyze RACE products and amplify LSINCT5 DNA. All primers used are listed in **Supplemental Table 2**. Briefly, 5' RACE and 3'RACE reverse transcription products were added to thin well micro-centrifuge tubes. Cycle proceeded as followed: Initial denaturation; 1 repeat at 94°C for 3 minutes, Amplification; 35 repeats at 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 5 minutes, Final extension 1 repeat at 72°C for 7 minutes.

DNA isolated from MCF7 cells as indicated in Qiagen DNeasy Blood and tissue kit (69504) protocol was used to amplify entire LSINCT5 transcript. Real-time PCR reaction for amplification of LSINCT5 using MCF7 DNA was done with the Roche FastStart High Fidelity PCR system (03553426001) following manufactures protocol.

Products were then loaded onto a 1% low melting agarose gel containing 1 ug/mL Fisher Bioreagents ethidium bromide (BP1302-10) and visualized on a UV transilluminator. All identified products were sent for sequencing for validation.

Quantitative real-time rt-PCR (qPCR). Primers for *LSINCT5*, *RACE* and all genes analyzed from microarray were designed using Primer3 (all primer sequences listed in **Sup. Table 2**). All cDNA samples were synthesized using the Superscript III RT-PCR system from Invitrogen (18080-044) using 150–200 ngs total RNA and Invitrogen OligodT primers (C110A) or Invitrogen Random Hexamers (N8080127). cDNA quantitation was then performed with specific primers using the Invitrogen Platinum SYBR green qPCR Supermix UDG with Rox (11744-500) and the ABI 7900HT Fast Real-time PCR system. Primers were optimized for qPCR with β -actin or U6 as a control gene and then with the transcript region of interest. When the optimal primer concentration produced a linear response to input cDNA concentration, RNA samples were analyzed in triplicate for each tested transcript. To normalize the expression levels (ΔC_T), the threshold cycle (C_T) for each transcript was subtracted from the C_T of the more abundantly-expressed control gene. Fold difference was calculated by subtracting normal ΔC_T tissue, cell line or NC ASO by tumor tissue, cell line or L5 ASO ΔC_T . Transcripts displaying consistently altered expression levels consistently in each of the triplicate experiments and were included for statistical analysis.

Localization. Nuclear and cytoplasmic fractions were isolated from BT474 cell line utilizing the Ambion PARIS (Protein and RNA Isolation System) kit (AM1921). Briefly, cells are washed with PBS and resuspended with cold fractionation buffer. Cells

are incubated on ice and centrifuged, and nuclear and cytoplasmic fractions are then aspirated off. Nuclear lysate is then lysed with disruption buffer. RNA is isolated from separate lysates by adding ethanol and filtering through a cartridge. Nuclear and cytoplasmic RNA (200 ng) is then converted to cDNA and analyzed by qPCR normalized with β -actin for different genes and LSINCT5 expression.

RNA polymerase inhibitors. BT474 cells were seeded in T75 flask and treated with 50 ug/ml Sigma Alpha Amanitin (A2263) or 65 uM Calbiochem RNA polymerase III inhibitor (557403) for 3 hours. RNA was then isolated from cells and qPCR was analyzed as stated previously.

Proliferation. Knock down of *LSINCT5* expression was conducted utilizing LSINCT5 antisense oligos (LSINCT5-1 ASO and LSINCT5-2 ASO) created using Integrated DNA technologies (IDT) AntiSense Design software using *LSINCT5* core sequence. Normal Control antisense oligo (NC ASO) was created by an IDT technical specialist. Sequences listed in **Supplemental Table 2**. OVCAR5 or MCF7 cells were seeded in a 96 well plate at 4×10^3 cells per well, transfected with both ASOs for LSINCT5 and NC ASO in MCF7 and OVCAR5 cancer cell lines for 24 or 48 hours using Invitrogen Lipofectamine 2000 Reagent (11668-027).

Proliferation was detected using Promega Cell Titer Blue reagent (G808B). OVCAR5 or MCF7 cells were seeded in a Thermoscientific 96 well white polymer pdl plate (152028) at 4×10^3 cells per well, transfected with Invitrogen Lipofectamine 2000 Reagent (11668-027) with 20 pM (OVCAR5) or 50 pM (MCF7) with both ASOs for LSINCT5 and NC ASO for 24 or 48 hours, cell titer reagent was added for 3 hours and incubated in 37° incubator, then analyzed on Tecan Spectrofluor Plus.

BrdU (5-bromo-2-deoxyuridine) analysis was conducted by staining cells with Thermoscientific BrdU (1860580). Briefly, MCF7 cells were transfected with 50 pM NC ASO or LSINCT5 ASO for 48 hours and fixed with 16% formaldehyde and permeabilized. Cells were then blocked and incubated with primary BrdU antibody for one hour at 37°. Secondary Dylight 488 goat anti-mouse antibody was added and finally nucleus was stained with DAPI. Cells were then evaluated using LSM510 laser microscope. Experiment was conducted twice in duplicate and quantification was done by counting 500 cells per experiment.

Affymetrix HG-U133 plus 2.0 gene expression array. Microarray experiment was performed using Affymetrix HG-U133 Plus 2.0 chips on 2 samples with knockdowns in MCF7 cells, one using LSINCT5 knockdown (LSINCT 5-1 ASO) and one normal scrambled oligonucleotide probe control (NC ASO) for forty-eight hours. A standard in-house MicroArray PreProcessing (MAPP) was used to preprocess the data. Differentially expressed genes between LSINCT 5 knockdown and control knockdown with fold changes greater than 1.5 were used for pathway analysis using MetaCore software (GeneGo Inc., St. Joseph, MI USA).

Acknowledgements

We acknowledge the Mayo Clinic Microarray Core Facility for assistance with gene expression array experiments. We would like

to thank Louis Maher III, Ph.D. and Rebecca R. Laborde, Ph.D. for revisions and comments on this manuscript. This work was funded by the Mayo Foundation and the Department of Defense Breast Cancer both as a Concept Award (W81XWH-07-1-0545), and recently as an Idea Award (NS 65007).

Note

Supplemental materials can be found at:
www.landesbioscience.com/journals/rnabiology/14800

AUTHOR: -Please edit references according to Landes guidelines. References in manuscript text should be in superscript, not brackets. See link below for examples: <http://www.landesbioscience.com/journalguidelines#refs> , please provide updated ref list as separate word doc. and provide when you return other corrections

References

- Craig Venter MDAJ, Myers EW, Li PW, Mural RJ, et al. Science 2001; 291:1304.
- Szymanski M. Ann NY Acad Sci 2006; 461.
- Consortium TEP. Nature 2007.
- Lamers SL, McGrath MS. Biosystems 2010.
- Filipowicz W. Nature 2008; 9:102.
- Garzon R, Fabbri M, Cimmino A, Calin GA, Croce CM. Trends Mol Med 2006; 12:580.
- Mercer TR, Mattick JS. Nat Rev Genet 2009; 10:155.
- Gabory A. Cytogenetic and Genome Research 2005.
- Chow JC, Ziesche SM, Brown CJ. Annu Rev Genetics Hu Genet 2005; 6:69.
- Rinn JL. Cell 2007; 129.
- Szegedi K, Nagy N, Németh IB, Bata-Csörgő Z, Kemény L, Dobozay A, Széll M. Exp Dermatol 2010; 19:269.
- Barciszewski SA. Gen Biology 2002; 3.
- Clemson CM, et al. Mol Cell 2009; 33:717.
- Fox AH, Bond CS, Lamond AI. Mol Biol Cell 2005; 16:5304.
- Fox AH, et al. Curr Biol 2002; 12:13.
- Ji P, et al. Oncogene 2003; 22:8031.
- Mallardo M. Journal of Experimental and clinical cancer research 2008.
- Guo F, Liu Y, Wang J, Li Y, Li G. Acta Biochem Biophys Sin 2010; 42:224.
- Geirsson A. J Heart Lung Transplant 2004; 9:1077.
- Silva JM, Pritchett JR, Halling ML, Tang H, Smith DL. Genomics 2010; 95:355.
- Ideue T, Kitao S, et al. RNA 2009.
- Chow JC, et al. Cell 141:956.
- Winter AG, Allison SJ, Tosh K, Scott PH, Spandidos DA, White RJ. Proc Natl Acad Sci USA 2000; 97:12619.
- Chen W, Brosius J, Tiedge H. J Pathol 1997; 183:345.
- Daly NL, Fairley JA, Gomez-Roman N, Morton JP, Graham SV, Spandidos DA, White RJ. Oncogene 2005; 24:880.
- Sonkoly E. J Bio Chem 2005; 25:21459.
- Hecht SS. Carcinogenesis 1988; 9:875.
- Lonardo F, Freemantle SJ, Ma Y, Memoli N, Sekula D, Knauth EA, et al. Clin Cancer Res 2002; 8:54.
- Chuang CH. Environ Mol Mutagen 2006; 47:73.
- Jorquera R, Schuller HM. Carcinogenesis 1994; 15:389.
- Abdel-Aziz HO, Tabuchi Y, Nomoto K, Murai Y, Tsuneyama K, Takano Y. J Cancer Res Clinical Onco 2007; 133:107.
- Sunwoo H, et al. Genome Res 2009; 19:347.
- Hirose ST. Genome Biol 2009; 10:227.
- Sasaki T, Ideue M, Sano T, Hirose MT. Proc Natl Acad Sci USA 2009; 106:2525.
- Brand S, et al. Exp Cell Res 2005; 310:117.
- Lee Y, et al. Oncogene 29:56.
- Hassan S, et al. Int J Cancer.
- Hamilton TC, Ozols RE. Semin Oncology 1984; 11:285.
- Conover CA, Hartmann LC, Bradley S, Stalboerger P, Klee GG, Kalli KR, Jenkins RB. Exp Cell Res 1998; 238:439.