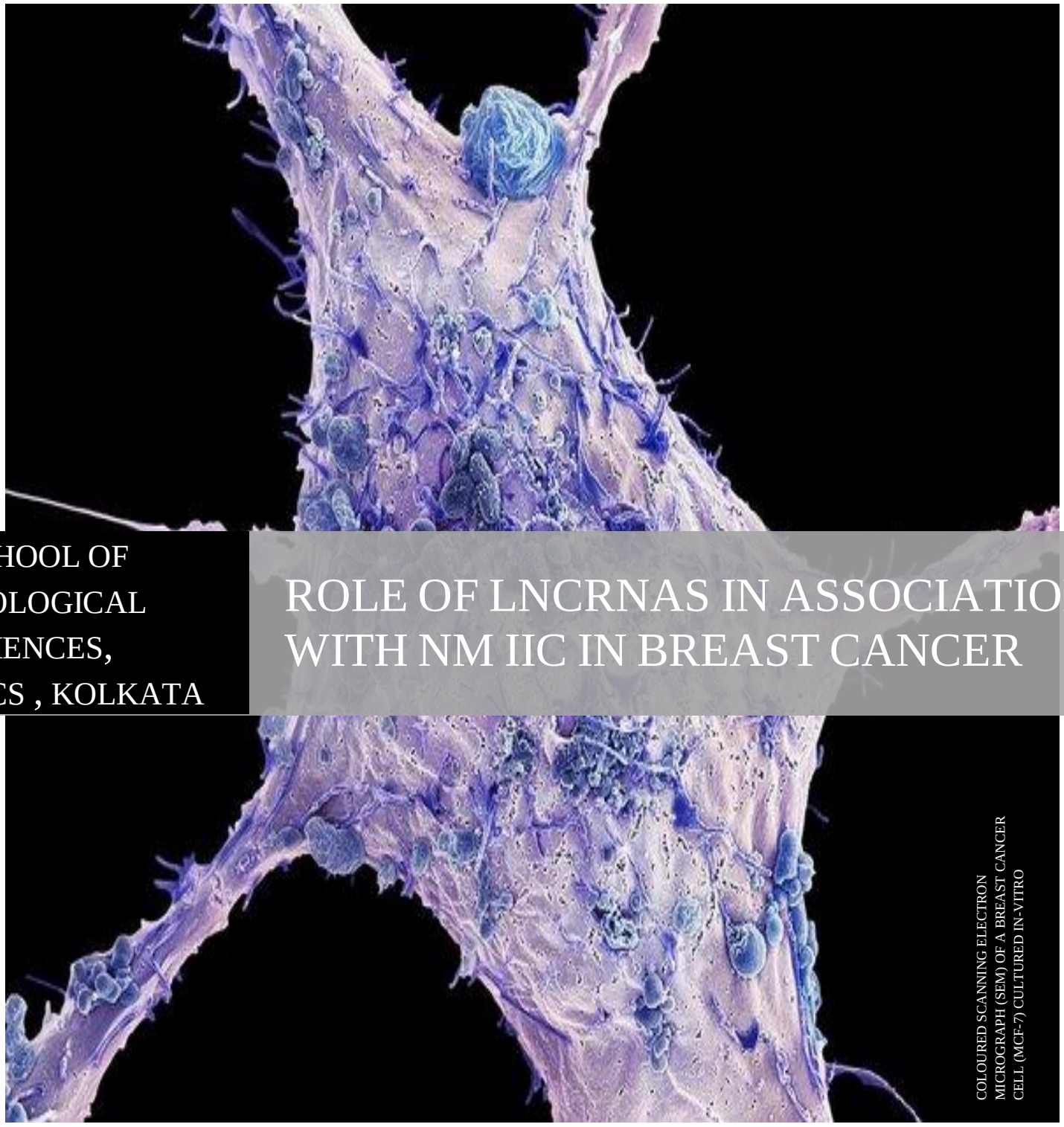

SCHOOL OF BIOLOGICAL SCIENCES, IACS , KOLKATA

ROLE OF LNCRNAS IN ASSOCIATION WITH NM IIC IN BREAST CANCER

COLOURED SCANNING ELECTRON MICROGRAPH (SEM) OF A BREAST CANCER CELL (MCF-7) CULTURED IN-VITRO

UNDER THE SUPERVISION OF PROF. SIDDHARTHA S. JANA, SCHOOL OF BIOLOGICAL SCIENCES, IACS, KOLKATA

BY

SHREYASHREE SARKAR, M.SC BIOTECHNOLOGY, JADAVPUR UNIVERSITY (2017-2019), ROLL NUMBER : 001720501017

इंडियन एसोसियेशन फर दि कल्टिवेशन आफ साइंस
ইণ্ডিয়ান এ্যাসোসিয়েশন ফর দি কালটিভেশন অব সায়েন্স
INDIAN ASSOCIATION FOR THE CULTIVATION OF SCIENCE

IACS/SBS/SSJ/2019-22

April 23, 2019

To Whom It May Concern

This is to certify that Ms. Shreyashree Sarkar, a student of Jadavpur University, worked under my guidance at Indian Association for the Cultivation of Science, Kolkata- 700032, India, from June 20, 2018 to April 23, 2019. She worked on the "Role of LncRNA in association with NM IIC in breast cancer". She could learn mammalian cell culture, total RNA isolation, siRNA transfection, PCR, western blot etc and also effectively analyse the experimental data on her own. She was very impressive, and maintained good laboratory practices during her stay in my laboratory at IACS. Please don't hesitate to contact me at bcssj@iacs.res.in if you have any query.

Yours truly,

Siddhartha S Jana

Dr. Siddhartha S Jana

Professor, School of Biological Chemistry

सिद्धार्थ एस. जाना, पीएचडी / Siddhartha S. Jana, PhD
प्राध्यापक / Professor
स्कूल ऑफ जैविक विज्ञान / School of Biological Sciences
आई. ए. सी. एस, जादवपुर, कोलकाता-32
IACS, Jadavpur, Kolkata-32

२ए एडं बी, राजा एस. सि. मल्लिक रोड, यादवपुर, कोलकत्ता-700 032, 2A & B, RAJA S. C. MULLICK ROAD, JADAVPUR, KOLKATA - 700 032
दुरभाष Phone : 2473-4971, 3372, 3073 फैक्स /FAX : (91) (33) 2473 2805, ग्राम / GRAM : इंडासन / INDASSON, यादवपुर / JADAVPUR

ACKNOWLEDGEMENT

This dissertation project would not have been possible without the guidance and motivation of my advisor, Prof. Siddhartha S. Jana and his lab which adapted the hypothesis of the regulatory roles of junk sequences of the genome, specifically the lncRNAs in the regulation of human breast cancer and their probable correlation with non muscle myosins. It is a result of his immense patience and encouragement that I have been able to grasp several concepts pertaining to our topic of research. I am thankful to Dr. Kavita Ghosal for teaching me complex molecular biological techniques and helping me strategize experiments with utmost care and dedication. The contribution of Debjit De, Research Scholar, Jadavpur University and Tilak Nayak, Research Scholar, Jadavpur University to this project is also highly acknowledged. Additionally, I would like to express my gratitude to the professors of Department of Life science and Biotechnology, Jadavpur University - Prof. Biswadip Das (H.O.D), Prof. Parimal Karmakar, Prof. Ratan Gachhui, Prof. Paltu Kumar Dhal, Prof. Arunima Sengupta and Prof. Satarupa Das for providing me with the opportunity to explore more of the scientific world with this summer project.

INDEX

LIST OF ABBREVIATIONS AND ACRONYMS USED

ABBREVIATION	DESCRIPTION
ATCC	American Type Culture Collection
BRCA	Breast Cancer Gene
DAPK3	Death Associated Protein Kinase 3
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphates
EDTA	Ethylenediaminetetraacetic Acid
EGF	Epidermal Growth Factor
ELC	Essential Light Chain
ER	Estrogen Receptor
EtBr	Ethidium Bromide
FBS	Fetal Bovine Serum
GBC	Gall Bladder Cancer
Her 2	Human Epidermal Growth Factor Receptor 2
HMM	Heavy Meromyosin
HOTAIR	Hox Transcript Antisense Ribonucleic Acid
L-15	Leibovitz 15 Media
LMM	Light Meromyosin
MBGW	Molecular Biology Grade Water
MEM-α	Minimum Essential Medium A
MHC	Myosin Heavy Chain
MLCK	Myosin Light Chain Kinase
NAT	Natural Antisense Transcripts
NCCS	National Centre For Cell Science
NM II	Non-Muscle Myosin II With Splice Variants A, B And C
PCAT19	Prostate Cancer Associated Transcript 19
PR	Progesterone Receptor
RLC	Regulatory Light Chain
RNA	Ribonucleic Acid
ROCK	Rho Associated Coiled Coil Containing Kinase
TAE	Tris Acetic EDTA
TBE	Tris Boric EDTA
ZIPK	Leucine Zipper Interacting Protein Kinase
siRNA	Small Interfering Ribonucleic Acid

ABSTRACT

Breast cancer is the most frequent malignancy in females. Due to its major impact on population, this disease represents a critical public health problem that requires further research at the molecular level in order to define its prognosis and specific treatment. [1] In recent years it has become clear that breast cancer does not represent a single disease but rather a number of molecularly-distinct tumors arising from the epithelial cells of the breast. [1] Breast cancer is highly heterogeneous encompassing a group of genetically and epigenetically distinct diseases exhibiting diverse clinical features, the most common and invasive of which is metastasis[2] which is more likely if the patient is suffering from triple negative breast cancer ($Her2^-$,ER^-, PR^-) [3]. NM IIs are non-muscle myosin protein responsible for cell shape, cell movement, transportation of substances in cells, cytokinesis, cell-cell adhesion, epithelial tissue morphogenesis and maintenance of the tensional homeostasis between cells in tissues. [4] Scientific breakthroughs in the research of non-coding sequences have shown that lncRNAs are capable of cis/trans regulation of several genes and this might also aid in the transformation of normal cells to cancerous cells. Once the functioning or localization of native NM IIs are disrupted, that the cell loses its optimal functioning and it is highly probable that the cell transform into tumorigenic or cancerous cells. [5, 6] Basic research is required to accomplish the task of finding correlation between lnc RNAs and NM-IIs in tumor progression in breast tissue. For the purpose, three types of human breast cell lines were procured from ATCC and NCCS, namely MCF10A (normal breast tissue), MCF7 (tumorigenic) and MDA-MB-231(metastatic) and one of the isoforms of NM IIs, NM IIC was considered for in-vitro analysis.

Two predicted lncRNAs were considered for the studying correlation between NMIIC and lnc RNAs.

One is HOTAIR, the one of the most popular lnc RNA for its role in cell migration, invasion and in EMT [75] especially in breast, found from published recognized literatures [76]. Another one is PCAT 19, a well known lncRNA for its dominant regulation over prostate cancer incidence. This PCAT19 is found to be expressing in different organs during tumorigenesis [39]. The position of the PCAT19 gene is sharing the same chromosomal location that is Chromosome 19 where MYH14 gene resides. Our interest of study is to unearth the possibility of companionship between the mentioned lncRNAs and NMIIC motor protein in human breast metastasis and this involves cell lines as they can be widely used in many aspects of laboratory research and, particularly, as in-vitro models in cancer research. The cell lines MCF-7 (tumorigenic) and MDA-MB-231 (metastatic) compared against healthy breast epithelial cell line-MCF-10A is used to understand and define whether two novel lncRNAs PCAT19 and HOTAIR, previously found to be involved in other forms of cancer[7,8] are involved in the development and progression of breast cancer as well.

INTRODUCTION

1. MYOSIN

The ability of cells to migrate is one of the crowning achievements of evolution. As cell movement is a mechanical work it requires two things ATP and the protein that can convert mechanical energy from ATP [9]. Cell can move by two types of process, one, by assembly and disassembly of microfilament and microtubules, and another mechanism requires special class of enzyme called motor proteins. Myosins are a super family of motor proteins involved in three basic functions namely, actin binding, and ATP hydrolysis and force transduction. [10] The myosin proteins are best known for their roles in muscle contraction and other motility processes in eukaryotes. All eukaryotic cells contain myosin isoforms. Myosin proteins are not restricted to muscle cells only and can be found in non-muscle cells as well.

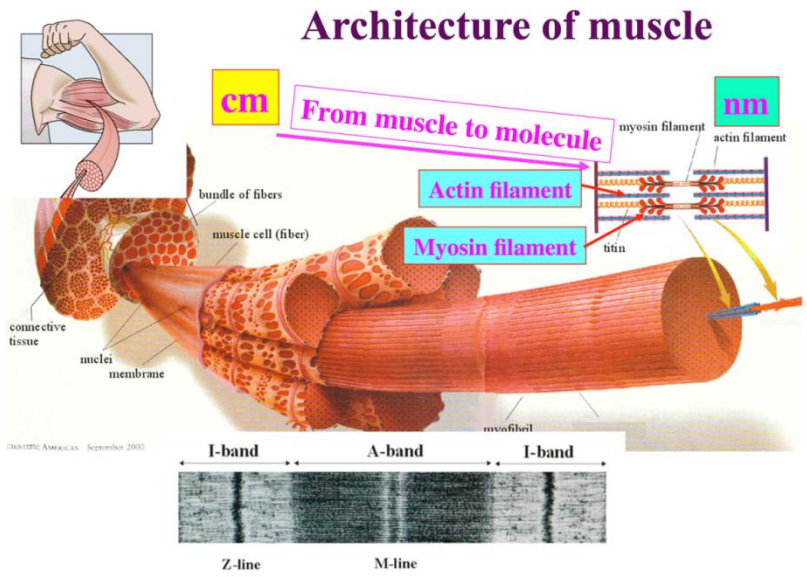

FIGURE 1: Architecture of Skeletal Muscle Formed By Regular Arrays Of Myosin And Actin Filaments

Virtually myosin proteins are found in all eukaryotic cells in differentially spliced isoforms. Some isoforms have specialized functions in certain cell types (such as muscle), while other isoforms are ubiquitous. The structure and function of myosin is globally conserved across species.

1.1 STRUCTURE AND FUNCTION

All myosin are composed of one or two heavy chains and several light chains. The heavy chains are organized into three structurally and functionally different domains: head neck and tail domain. The structure and function of the head neck and tail domain are as follows-

a. **Head**: The Head domain is a bulbous projection that binds ATP to the filamentous actin and generates the force to walk along the actin filament from the negative end to the positive end. This is the most conserved region among the myosin.
b. **Neck**: The α-helical neck domain is associated with the light chain which regulates the activity of the head domain.
c. **Tail**: The tail domain contains the binding sites that determine the specific activities of a particular myosin.

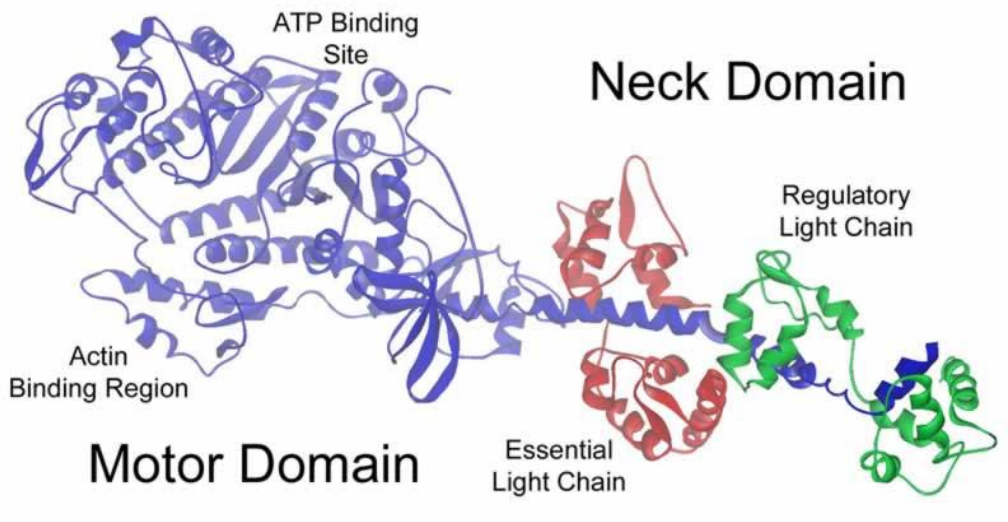

FIGURE 2: Structural Domains of Myosin

Thirteen members of the myosin gene family have been identified by genomic analysis. Myosins are mainly of two types: (a) **Conventional Myosin** and (b) **Unconventional Myosin**. Myosin II belongs to conventional myosin. Among the other myosin, conventional myosin has a characteristic to self-assemble into the bipolar filaments through tail-tail interaction. In vertebrates, there are three types of myosin found, myosin I, myosin II, and myosin V. Myosin I and Myosin II are the most abundant and studied types whereas Myosin V, a less common isoform has also been isolated and characterized. Myosin II powers muscle contraction and cytokinesis, whereas Myosin I and V are involved in cytoskeleton membrane interactions such as the transport of membrane vesicles. Myosin II and Myosin V are dimers in which α-helical sequences in the tail of each heavy chain associated to form a rod like coiled coil structure. Because Myosin I heavy chain lacks this α-

helical sequence, the molecule is a monomer. The three myosin differ in the number and type of light chains bound in the neck region. Myosin IIs are can be further classified into three types:

- SMOOTH MUSCLE MYOSIN
- SKELETAL AND CARDIAC MUSCLE MYOSIN
- NON-MUSCLE MYOSIN (NM)

Of the above classifications of Myosin II, Non-muscle myosin II (NM II) is the area of concentration in this project work. Phylogenetically, the NM IIs are closer to smooth-muscle myosin than they are to the sarcomeric myosin. However, the name 'nonmuscle myosin' remains a misnomer in that these non-muscle myosin II isoforms are also present in all muscle cells (skeletal, cardiac and smooth), though in significantly smaller quantities than the muscle myosin IIs. In mammals, three different isoforms of nonmuscle myosin II, II-A, II-B and II-C, are widely distributed throughout the entire organism. These isoforms generate due to the difference in the length of the heavy chain of the myosin protein. While a few cells contain a single isoform, most contain more than one, including isoforms generated by alternative splicing. In humans, these isoforms are encoded by three different genes, myh9 (IIA), myh10 (II-B) and MYH14 (II-C), present on three different chromosomes [11, 12]. These proteins play a role in many fundamental cellular and developmental processes such as cell-cell adhesion, cell migration and cytokinesis.

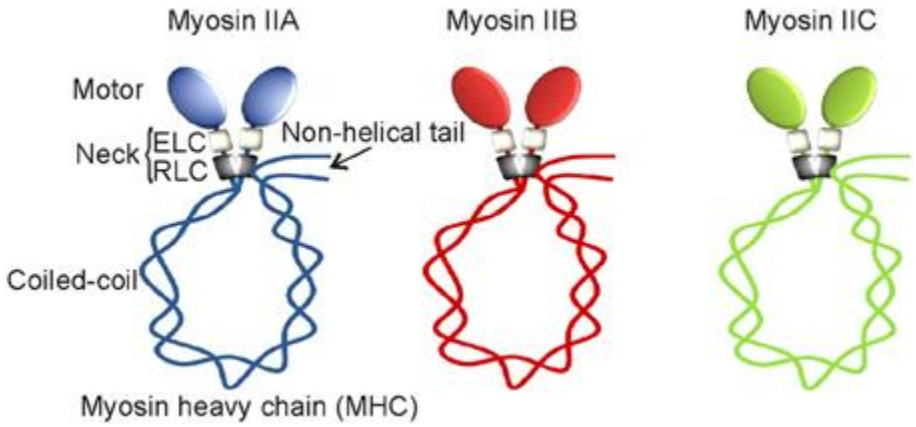

FIGURE 3: Heavy and Light Chains Of Different Isoforms Of Myosin

Although all three isoforms share a number of biochemical and structural properties, there are also important differences among them that are being investigated at both the cellular level as well as in whole animals. Small interfering RNAs specific for each of the isoforms as well as the relatively specific inhibitor of myosin MgATPase activity, blebbistatin, are two important new tools helping to elucidate the function of nonmuscle myosin II. The use of homologous recombination to generate mice that have been ablated for, or have markedly decreased amounts of each isoform, or have point mutations in the various isoforms, is also providing new information about the role of these proteins in vivo [12] The purpose of the present chapter is to review the

properties of nonmuscle myosin II and its signaling pathways and to provide examples of its functions in cells from a number of species, as well as model animals.

1.2 CROSS BRIDGE CYCLE

Multiple myosin II molecules generate force in skeletal muscle through a power stroke mechanism fuelled by the energy released from ATP hydrolysis. [13] The power stroke occurs at the release of phosphate from the myosin molecule after the ATP hydrolysis while myosin is tightly bound to actin. The effect of this release is a conformational change in the molecule that pulls against the actin. The release of the ADP molecule leads to the so-called rigor state of myosin [14]. The binding of a new ATP molecule will release myosin from actin. ATP hydrolysis within the myosin will cause it to bind to actin again to repeat the cycle. [15] The combined effect of the myriad power strokes causes the muscle to contract with the help of this cross bridge cycle.

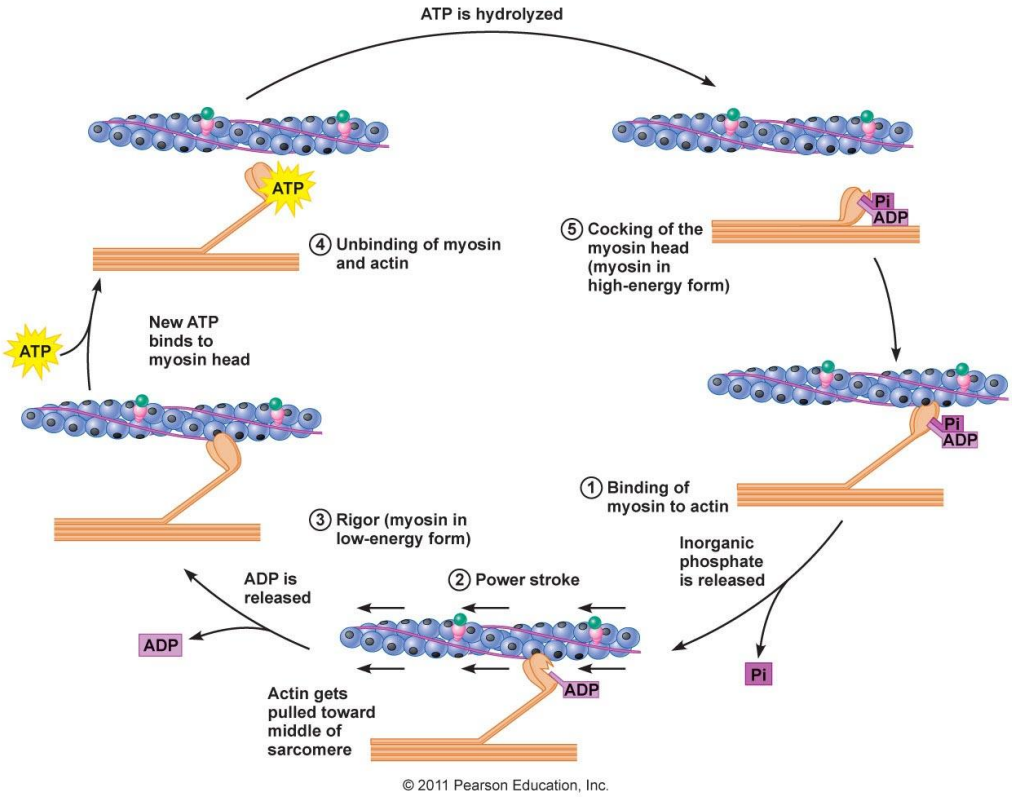

FIGURE 4: Cross Bridge Cycle/Power Stroke to Generate Muscle Movement

1.3 STRUCTURAL BACKGROUND OF NM II

The class II myosins, which were the first to be discovered and are therefore often referred to as "conventional myosin," are filament-forming proteins and include those responsible for skeletal, smooth, and cardiac muscle contraction. In common with all members of the myosin II class, nonmuscle myosin II (NM II) is a hexamer formed by dimerization of two heavy chains, each ~2,000 amino acids long, and each heavy chain binds two light chains. [16]

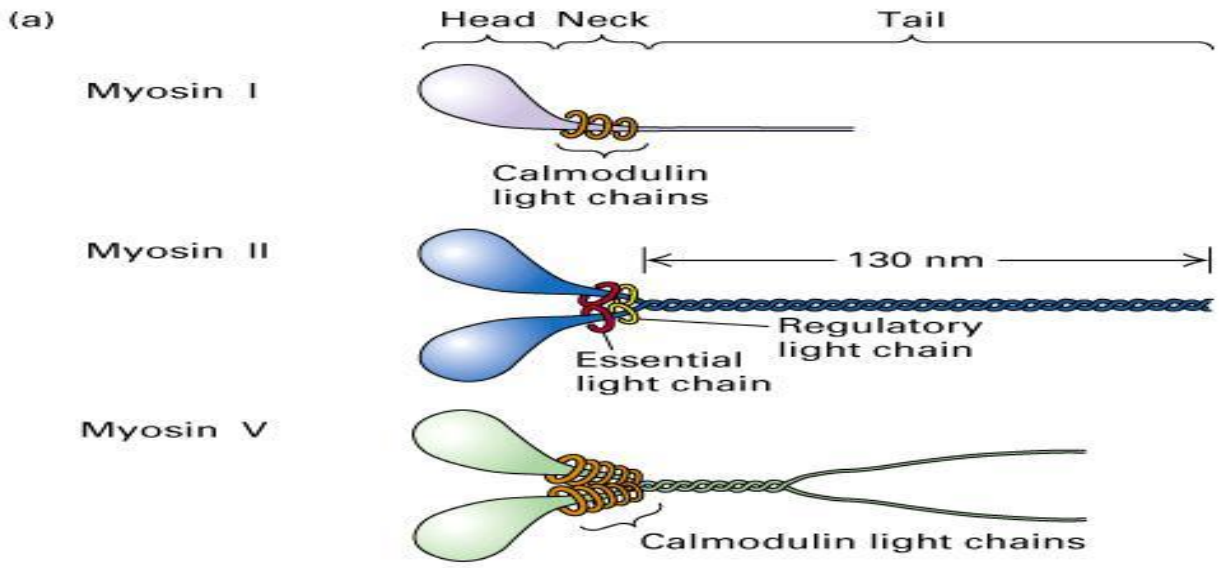

FIGURE 5: Types Of Myosin And Their Structure

NM II has a very asymmetric shape with a globular motor head domain at the N-terminal (containing a binding site for both ATP and ACTIN) [17], followed by a neck region (containing two regulatory light chain of 20KDa that regulate NM II activity and two essential light chain of 17KDa which mainly regulate the heavy chain structure), an α-helical coiled coil domain (containing two heavy chain of 230KDa) and a non-helical tail domain. The heavy chain begins with an N-terminal motor domain that is responsible for ATP hydrolysis and is the source of force generation. Following this is a lever arm region that binds the light chains. Proximal to the motor domain is the essential light chain (ELC) that acts to structurally stabilize the lever arm.[21] The remainder of the lever is stabilized by the regulatory light chain (RLC), which plays a role in regulating the enzymatic activity of the myosin in a phosphorylation-dependent manner. C-terminal to the lever arm is a long (~1,100 amino acids) α-helical region responsible for dimerization of the heavy chains to form the characteristic coiled coil tails of myosin II molecules. The first 800 amino acids constitute head domain and this is the conserved

domain of myosin heavy chain. Then the neck region which is the binding side for RLC and ELC has two very specific amino acids **isoleucine** and **glutamine**. Light chains mainly bind to IQ motifs and it is in these regions

That the above two amino acids involved. Deletion of IQ 1 and IQ 2 motifs then binding of light chain cannot be possible. A non-helical region at the tip of the tail terminates the heavy chain. [33]

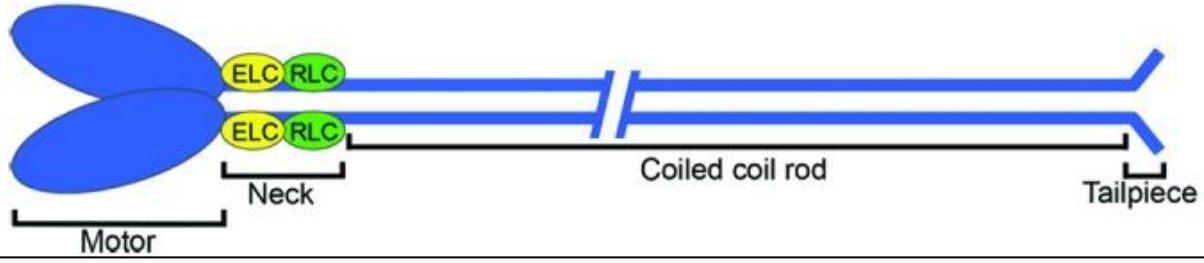

FIGURE 6: Domain organization of myosin-II. Each myosin-II heavy chain is composed of an N-terminal motor domain, a neck domain that binds one essential light chain (yellow) and one regulatory light chain (green), a coiled coil rod domain and a C-terminal tailpiece. The head domain contains the ATP binding site and the actin binding domain. The two myosin-II heavy chains dimerize via interactions between the rod domains, which form an α-helical coiled coil. This hexameric assembly is referred to as the myosin-II monomer. [33]

Controlled proteolysis of NM II produces two fragments Heavy Meromyosin (HMM) and Light Meromyosin (LMM), further digestion of HMM formed Sub fragment 1 (S1) and Sub fragment 2 (S2). Then again digestion of S1 produces three fragments with 25KDa, 50KDa and 20KDa. The surface loop1 present between 25KDa and 50KDa whereas loop2 present between 50KDa and 20KDa. [18]

1.4 ALTERNATIVE SPLICING OF NM IIS

Alternative splicing, or differential splicing, is a regulated process during gene expression that results in a single gene coding for multiple proteins. The production of alternatively spliced mRNAs is regulated by a system of trans-acting proteins that bind to cis-acting sites on the primary transcript itself. Such proteins include splicing activators that promote the usage of a particular splice site, and splicing repressors that reduce the usage of a particular site. Mechanisms of alternative splicing are highly variable, and new examples are constantly being found, particularly through the use of high-throughput techniques. MHC II isoform diversity is generated by multiple genes as well as by alternative splicing of pre-mRNA. The three isoforms of NM IIs undergo alternative splicing at loop 1 and loop 2 region of their own heavy chain and form theirs isoforms. NMHC II-B and NMHC IIB and NMHC IIC have spliced variants NM IIB1, NM IIB2 and NM IIC1, NM IIC2, and NM IIC1C2 respectively. [30] B1 and C1 exons encoding 10 and 8 amino acids respectively are inserted into loop 1 at amino acids 212 and 227 near ATP binding site of their corresponding heavy chain. [31, 32] The spliced variant of NM IIA isoform is NM IIA2; in this case the insertion of amino acids occurs at loop 2 regions. A new RNA binding protein family, the Rbfox family, plays a critical role for neuron-specific alternative splicing of NMHC II-B pre-mRNA synergistically acting with PSF.

Isoforms of non-muscle myosin II-C (NM II-C), NM II-C2 is generated by alternative splicing of an exon, C2, encoding 41 amino acids in mice (33 in humans). The 41 amino acids are inserted into loop 2 of the NM II-C heavy chain within the actin binding region.

1.5 REGULATION OF NON- MUSCLE MYOSIN II

Myosin II exhibits enhanced binding to **actin filaments** under tension [15] and will accumulate to regions of high stress within the cell. This ability to detect and respond to forces is known as mechanosensation. Mechanosensing provides an important mechanism for directing localization of myosin II and is fundamental to a wide range of cellular and tissue functions [34]. Besides, mechanosensing also plays a major role in cleavage furrow concentration of myosin II. [35] Actin mediators, being allosteric, are ideal mediators of cooperative interactions of NM IIC as described above. [13, 14, 15]

Under normal condition NM II present in compact form which folded structure sediments at 10S, that is called 10S form. They are unable to form bipolar filaments. After **phosphorylation at RLC**, it exists in an elongated conformation due to C-terminal tail detachment, this folded structure sediments at 6S and it is called 6S form that is competent form and then assemble into bipolar filaments. Although 10S and 6S both conformation remains in equilibrium with each other and the equilibrium of both them depends upon the solvent condition. RLC, in 10S form, is associated with neck region of NMHC (according to the left side picture) which undergoes reversible phosphorylation on its S1, S2, S3 , S19, T18 amino acids and thereby myosin motor protein are turn on the activity of myosin protein and on another side dephosphorylation of S19 and T18 or phosphorylation at S1, S2, S3, T9 causes inactivation of the myosin motor protein.RLC reversible phosphorylation is tightly regulated by both myosin specific phosphatase [19] and wide variety of kinase enzymes like myosin light chain kinase [MLCK], Rho associated coiled coil containing kinase [ROCK][20], leucine zipper interacting protein kinase [ZIPK], death associated protein kinase 3[DAPK3] etc., and all of this kinases are involved in phosphorylation of S19 and T18 amino acid residues of NMHC which turn on myosin. But another type of kinase Protein Kinase C (PKC) causes phosphorylation at S1, S2, S3 [19, 21, 22] T9 residues for inactivation of myosin during cell division process. Although these states are very difficult to visualize in live cells, some recent reports suggest that 10S-6S transition occur in live cells [23].

Tropomyosins can regulate actomyosin by inhibiting or activating actomyosin MgATPase activity and motility depending on the myosin and its isoforms. Tropomyosins activated, inhibited, or had no effect on motility depending on the myosin, indicating that the myosin isoform is the primary determinant of the isoform-specific effect of tropomyosin on actomyosin regulation. Activation of motility of nonmuscle tropomyosin–actin filaments by NMII myosin correlates with an increased V_{max} of the myosin MgATPase, implying a direct effect on the myosin MgATPase, in contrast to the skeletal tropomyosin–actin filament that has no effect on the V_{max} or maximal filament velocity. [24, 25]

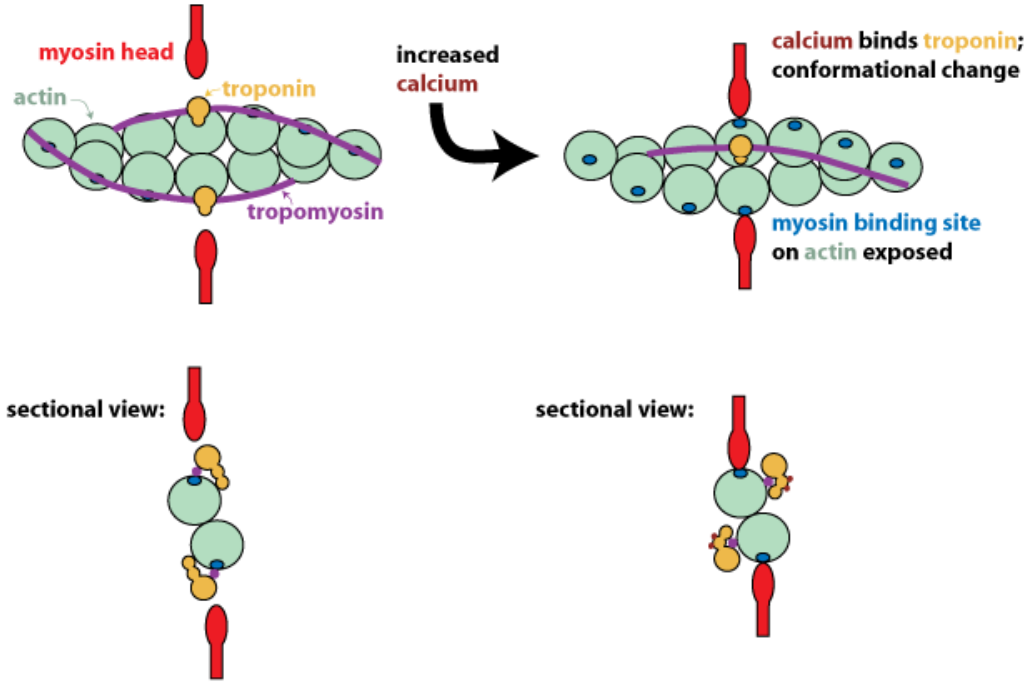

FIGURE 7: Ca^{++} Regulation Of Myosin Mediated By Tropomyosin

Interaction between actin and myosin in non-muscle cells and in vertebrate smooth muscle is regulated by the level of phosphorylation of the 20,000-molecular weight (MW) light chain. In the absence of calcium, this light chain is not phosphorylated and the myosin cannot interact with actin. Calcium activates a specific calmodulin-dependent kinase which phosphorylates the light chain, initiating actin-myosin interaction. Although most studies on the role of phosphorylation have concentration on the regulation of actin-activated myosin Mg-ATPase activity, phosphorylation of the light chain also seems to control the assembly of smooth muscle myosin into filaments. [21]

2. Lnc RNA

In contrast to a small proportion of the mammalian genome (e.g., human, mouse) that are transcribed into mRNAs, the vast majority of the genome is transcribed into what was previously regarded as "dark matter"—non-coding RNAs (ncRNAs) that do not encode information about proteins.[26]Among these ncRNAs, long ncRNAs (lncRNAs) represent the most prevalent and functionally diverse class.[27,28] Long non-coding RNAs (lncRNAs) are a class of transcribed RNA molecules with a length of more than 200 nucleotides that do not encode proteins (or lack > 100 amino acid open reading frame). lncRNAs are thought to encompass nearly 30,000 different transcripts in humans, hence lncRNA transcripts account for the major part of the non-coding transcriptome. LncRNA discovery is still at a preliminary stage. lncRNAs can be transcribed as whole or partial natural antisense transcripts (NAT) to coding genes, or located between genes or within introns. Some lncRNAs originate from pseudo genes [29]. Lnc RNA genes are tissue specific although cell to cell variation in the expression levels and type is found which suggests the idea of alternative splicing.[30, 31] Lnc RNAs are poorly conserved between species. In quite a few cases it has been found that truncated versions of Lnc RNAs are produced which, after translation become highly unstable and non- functional.

2.1 CLASSIFICATION

Lnc RNAs are classified into the following types:

- Intergenic LncRNA – This type of lncRNA is transcribed from the intergenic regions of the DNA
- Intragenic LncRNA – This type of lncRNA is transcribed entirely from introns. Intergenic lncRNAs and intronic lncRNAs are most likely regulated through different transcription activation mechanisms[52] and may have different poly (A) modifications and manifest activities in different cellular locations.[53]
- Sense LncRNA – This type of lncRNA is transcribed from the sense strand of protein-coding genes, containing exons from protein-coding genes. They may overlap with part of protein-coding genes, or cover the entire sequence of a protein-coding gene
- Antisense LncRNA – This type of lncRNA is transcribed from the antisense strand of protein-coding genes. According to GENCODE (a database of manually curated lncRNAs) annotation,[51] antisense lncRNAs may appear in three scenarios: (1) transcripts from the antisense strand of protein-coding genes overlap an exon of a sense gene through lncRNAs' exons, (2) transcripts from the intron of a sense gene do not have exon-exon overlap with this sense gene and (3) transcripts cover the entire sequence of a sense gene through an intron

LncRNAs can be further classified more stringently. Variants include: divergent lncRNAs, nested antisense lncRNAs, terminal antisense lncRNAs, small RNA containing or stand alone lncRNAs, pseudogenic lncRNAs, promoter associated lncRNAs etc.

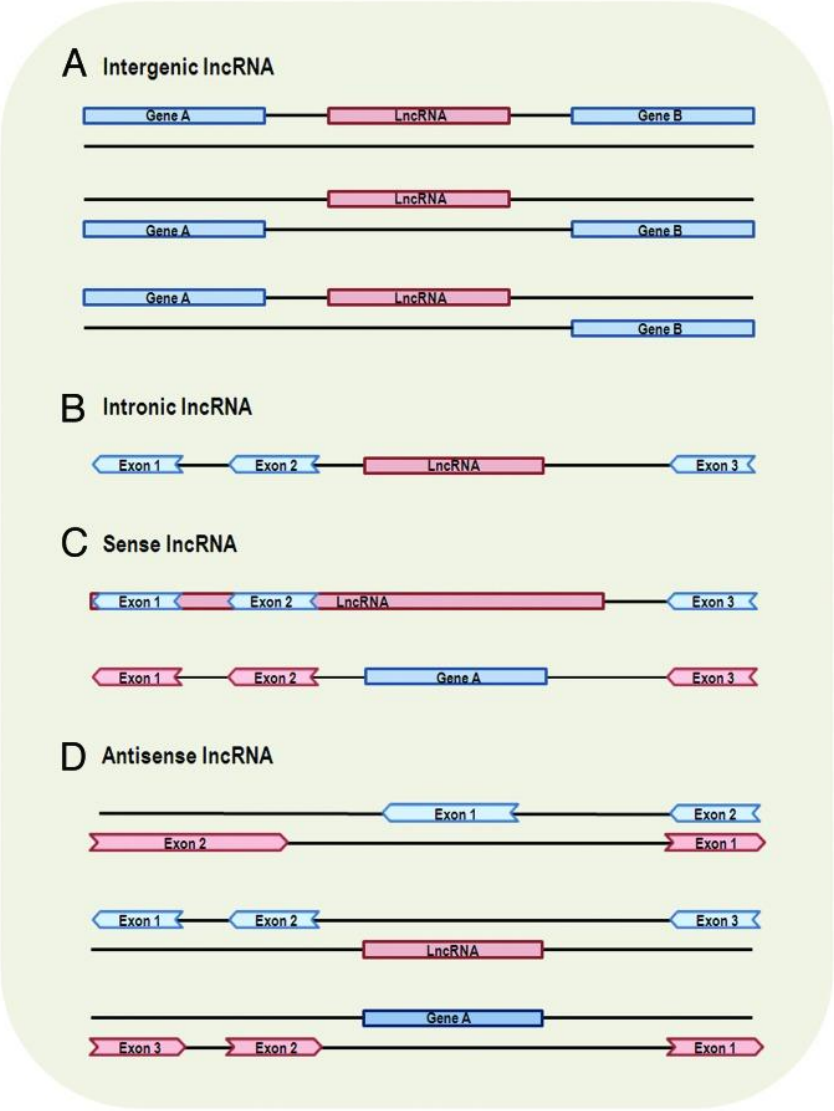

FIGURE 8: Types of LncRNA

2.2 FUNCTIONAL MECHANISMS OF LncRNA.

LncRNAs are predominately localized in nucleus and chromatin, [51] suggesting that lncRNAs may have a significant impact on DNA sequences. Also, a large proportion of lncRNAs are involved in transcriptional regulation (~42% of the 182 assessed entries according to lncRNA database. Therefore, it is meaningful to classify lncRNAs based on their effects exerted on DNA sequences:

- Cis-lncRNAs (cis-acting lncRNAs) that regulate the expression of genes in close genomic proximity e.g., DHFR (dihydrofolate reductase) upstream transcripts,[53,54] Xist (X inactive-specific transcript)[55]

- Trans-lncRNAs (trans-acting lncRNAs) that regulate the expression of distant genes. E.g.: HOTAIR (HOX antisense intergenic RNA), a ~2.2 kb lncRNA that is transcribed from the HOXC (homeobox C cluster) gene locus in chromosome 12, can be transported by the Suz-Twelve protein to regulate the homologous target sites at HOXD (homeobox D cluster) gene locus in chromosome 2. unlike cis-lncRNAs, trans-lncRNAs may function independently of sequence complementary to target gene locus. In addition to chromatin modification complexes, [56, 57] they may bind to transcription elongation factors [57] or RNA polymerases [58] to affect transcription.[59]

LncRNAs are involved in a wide variety of cellular molecular functions. According to their mechanisms of functioning, lncRNAs roughly fall into three groups that affect transcriptional regulation, translational regulation and other functions

- Transcriptional regulation - lncRNAs responsible for transcription regulation can be sub-divided according to the mechanism of their functioning: (1) transcriptional interference and (2) chromatin remodeling Besides, there are other related functional mechanisms, for example, the regulation effect: a set of lncRNAs transcribed from enhancers are termed eRNAs (enhancer RNAs) as they positively regulate genes' transcription, for instance, ncRNA-a1 (activating long ncRNA 1).

- Translational regulation - There are two common translational regulation mechanisms that lncRNAs get involved in, namely, splicing regulation and translational control. LncRNAs that influence mRNA splicing may function through binding to [61] or modulating. [60] splicing factors, or directly hybridizing with mRNA sequences to block splicing. Malat1 (metastasis-associated lung adenocarcinoma transcript 1), a ~7 kb lncRNA, can bind to SR splicing factor [serine–arginine (SR)-rich splicing factor] and regulate its distribution in nuclear speckle domains. LncRNAs that participate in translational control may function through binding to translation factors [62, 63] or ribosome.

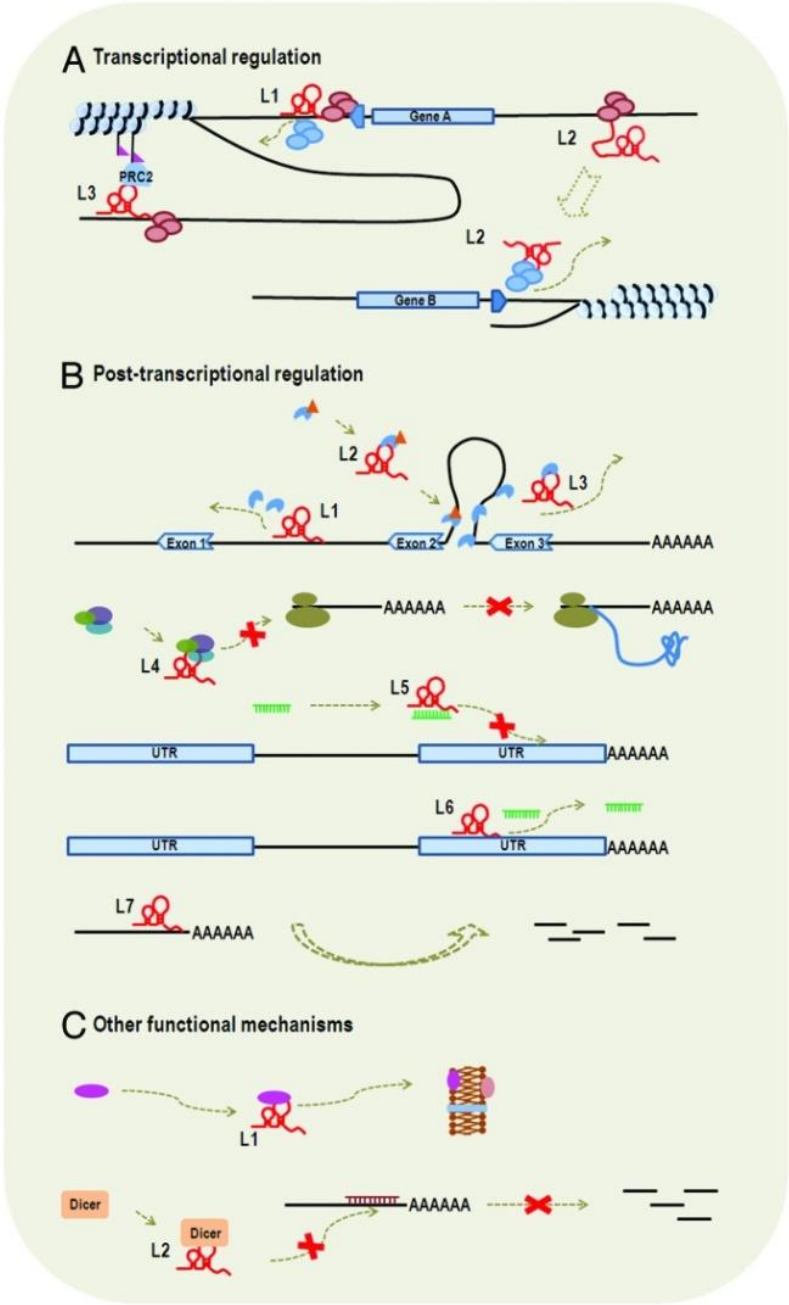

FIGURE 9: Functional mechanisms of lncRNAs. LncRNA is represented by a letter "L" and a number appended. (**A**) Transcriptional regulation: We listed examples of *cis*-lncRNA (L1 and L3) and *trans*-lncRNA (L2). L1 is transcribed from the promoter region of gene A and its binding to promoter of gene A blocks the binding of transcription factors, thus affecting transcription initiation of gene A. L3 functions to modify chromatin protein in its vicinity through recruiting the complex of PRC2. L2 influences transcription of gene B from a distant region through interaction with transcription factor or RNA polymerase. Therefore, L1 and L2 also function through transcriptional interference, whereas L3 functions through chromatin modification. (**B**) Post-transcriptional regulation: L1, L2 and L3 all influence gene splicing. Specifically, L1 binds to intronic area to inhibit binding of splicing factor, L2 functions to modulate the pool of modified (such as phosphorylation) splicing factor and L3 binds to splicing factor to block spliceosomal complex formation. L4 interacts with translational factors to inhibit translation. L5 and L6 are two examples of ceRNAs, which interact directly or indirectly with miRNAs. L5 binds to miRNA and, thus, inhibits the binding of miRNA to the 3′ UTR of target mRNA. L6 binds to the 3′ UTR of target mRNA, which also blocks the binding of miRNA to the target gene. L7 serves as natural antisense inhibitor to promote degradation of mRNA. (**C**) Other functional mechanisms. L1 is involved in protein transportation and L2 binds to Dicer to influence RNA interference.

LncRNAs may also function as natural antisense inhibitors to promote degradation of mRNA. Moreover, there are many lncRNAs that interact directly or indirectly with miRNAs to stabilize target mRNAs. lncRNAs may function through other mechanisms, such as protein localization,[64, 65] telomere replication,[66] RNA

interference, epigenetic regulation [66] beyond transcription and translation regulation, and other yet to be known mechanisms. In addition, lncRNAs may be involved in RNA interference by regulation of Dicer1.

2.3 TARGETING MECHANISMS OF LncRNA

- SIGNAL- To show cell type-specific expression and respond to diverse stimuli and functions as molecular signal. E.g. -COLDAIR
- DECOY- To bind and titrate away a protein target, but does not exert any additional functions, such as DHFR upstream transcripts.
- GUIDE- To bind proteins and then direct the localization of ribonucleoprotein complex to specific targets, such as HOTAIR;[67,68]
- SCAFFOLD- To serve as central platforms to bring together multiple proteins to form ribonucleoprotein complexes, such as HOTAIR
- ENHANCER- To control higher order chromosomal looping in an enhancer like model.

Alternative splicing of lncRNAs regulate their expression variation in various tissues and this is the key feature why lncRNAs are adept in intracellular and intercellular regulation. The up regulation or down regulation of tissue specific lncRNAs in close association with the NMIIs, specifically NM IIC (which is found to be associated with luminal breast cancer cells) interferes with the cancer cell adhesion and tensional homeostasis which leads to formation of invadopodia and lamellipodia, EMT, invasion into other healthy tissues, aggression, secondary tumor formation and hence metastasis.

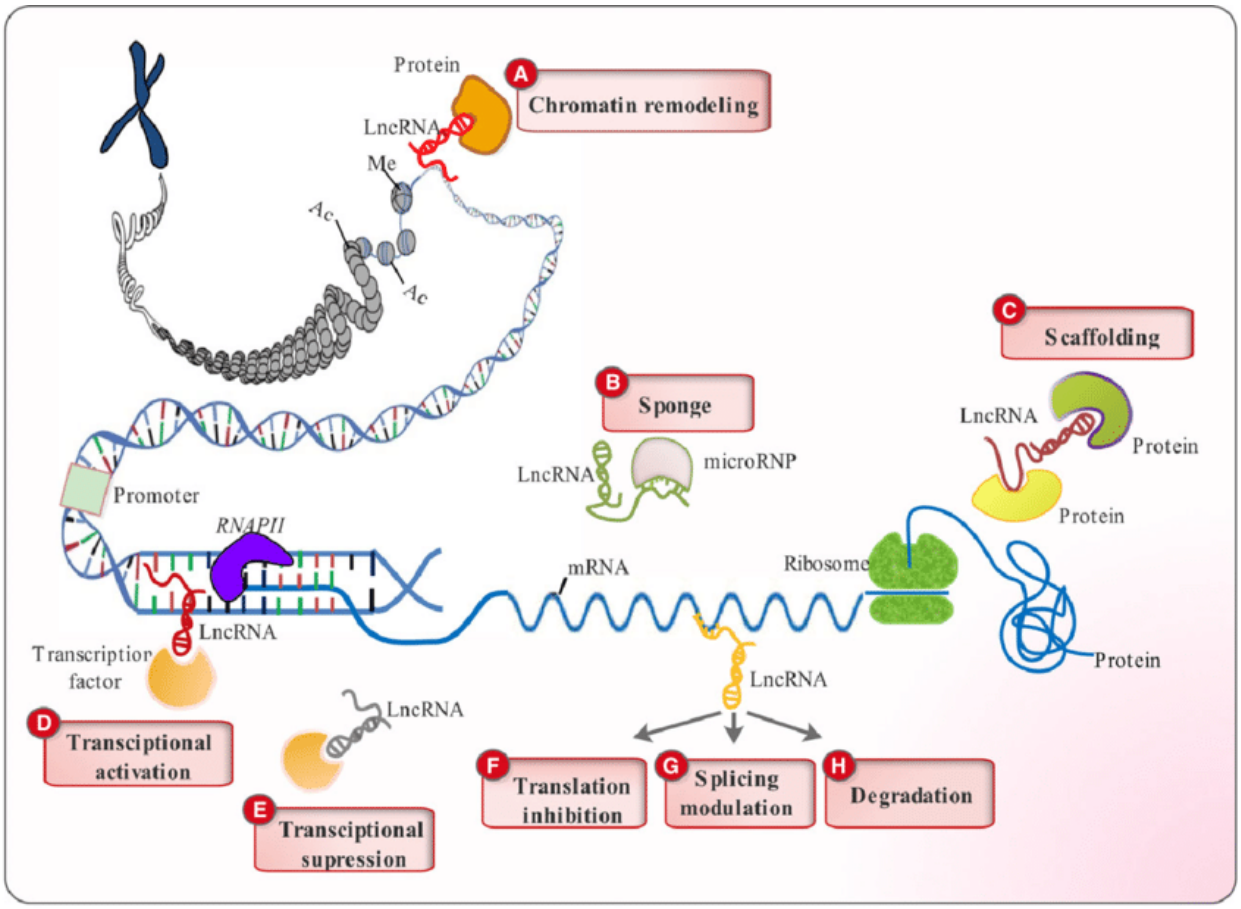

FIGURE 10: Targeting Mechanisms of LncRNA [37]

This project concentrates on the effect of two lncRNAs previously found associated with cancer namely HOTAIR and PCAT19 in human breast cancer.

- HOTAIR-
    a. LOCATION AND SIZE: HOTAIR is a human gene located on chromosome 12. The HOTAIR gene contains 6,232 bp and encodes 2.2 kb long non-coding RNA molecule, which controls gene expression.[40] Its source DNA is located within a HOXC gene cluster. [41] It is shuttled from chromosome 12 to chromosome 2 by the Suz-Twelve protein from where it can trans regulate chromatin state of a cell.[40,41,42]

    b. MODE OF ACTION: HOTAIR lncRNA has been found to be overexpressed metastatic cancer of breast. HOTAIR also mediates over expression of the HER2 oncogene . In cells, especially those that over express PRC2. The 5' end of HOTAIR lncRNA interacts with Polycomb Repressive Complex 2 (PRC2) and causes histone H_3 lysine-27 trimethylation which eventually leads to silencing of genes.[43]Therefore, In cells, especially those that over express PRC2, the prevention of HOTAIR expression leads to a reduction in invasive potential of that cell. [40] HOTAIR also plays a crucial role in EMT and in maintenance of

cancer stem cells. It has been found to be associated with other types of cancer such as gastric cancer [44], colorectal cancer [45], lung cancer [46] and prostate cancer [47].

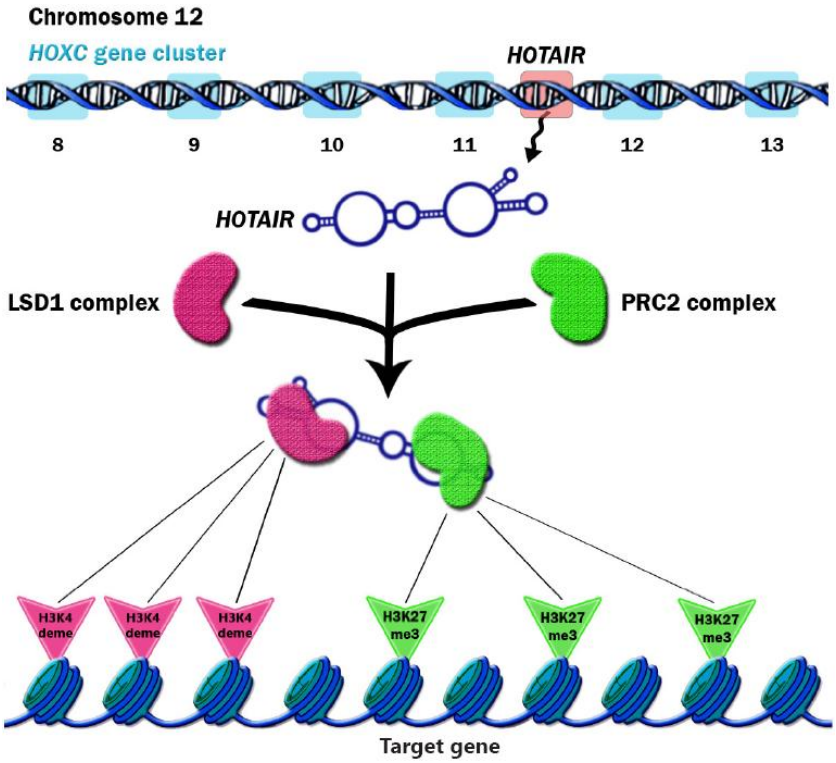

FIGURE 11: HOTAIR gene is located on chromosome 12 inside the HOXC locus, specifically between HOXC11 and HOXC12. After the expression of HOTAIR, this lncRNA recruits PRC2 and LSD1 complexes and thus functions as a bridge. HOTAIR directs these complexes to their target genes and as a result regulates the trimethylation of H3K27 and demethylation of H3K4 at targeted genes. H3k27Me3 and h3k4DeMe refer to the trimethylation of histone H3 at lysine-27 and the demethylation of histone H3 at lysine 4, respectively.

- PCAT19-
    a. LOCATION AND SIZE: PCAT19 is a 40887 bp linear DNA sequence present in the human chromosome 19. This sequence constitutes of 1 gene allele and 7 splice variants of intergenic lncRNA.

    b. MODE OF ACTION: PCAT19 is a relatively new lncRNA discovered, and hence researchers are still skeptical about its mode of function. In one experiment, the expression levels of PCAT19 was found to be increased in laryngeal tumor tissues and associated with decreased overall survival as compared to normal laryngeal tissues. Using laryngeal cancer cells lines, it was demonstrated that knockdown of PCAT19 decreased the cell proliferation, increased the mitochondrial respiration, and inhibited the glycolysis. Finally, researchers have shown that the PCAT19 knockdown decreased the tumor growth in vivo in laryngeal cancer. [48] Similarly, the risk variant of prostate cancer has been found to be associated with decreased and increased levels of PCAT19-short isoform and PCAT19-long isoform, respectively. [49] In another experiment comparative study of differential expression of lncRNAs with mRNAs in normal and tumorigenic gallbladder tissues revealed that a total of 89 up regulated (13 lncRNAs and 76 mRNAs) and 261 down regulated transcripts (27 lncRNAs and 234 mRNAs) were identified in the GBC tissues one of which is PCAT19. Thus in conclusion, it can be said that lncRNA PCAT19 promotes cell proliferation and tumorigenesis in various tissues. [50] This may provide a foundation for the development of credible therapeutic approaches with lincRNA PCAT19 as target.

3. BREAST CANCER CELL LINES

Breast cancer cell lines have been widely used for breast cancer modeling which encompasses a panel of diseases with distinct phenotypical associations. Though cell lines provide unlimited homogenous materials for tumor studies and are relatively easy to culture, they are known to accumulate mutations during the initial establishment and subsequent series of cultivations. Thus, whether breast cancer cell line heterogeneity reflects that of carcinoma remains an important issue to be resolved before drawing any reliable conclusion at the tumor level using cell lines. After combing through several breast cancer cell lines, analyzing the molecular features, genetically and epigenetically, of each subtype three cell lines were selected:

3.1 MCF-10A

The MCF10A human mammary epithelial cell line is a widely used in vitro model for studying normal breast cell function and transformation.

- CULTURE- MCF10A cells are cultured in DMEM/Ham F-12 media (GIBCO-Invitrogen) supplemented with 100 ng/ml cholera toxin, 20 ng/ml epidermal growth factor (EGF), 0.01 mg/ml insulin, 500 ng/ml hydrocortisone, and 5% chelex-treated horse serum.. MCF10A cells were subjected to no more than eight passages in culture when used in experiments.
- MORPHOLOGY- The cell takes up within 24-36 hours of seeding and is visible in a spheroid shape. Besides expression of stem cell/progenitor markers like EpCAM+/CD49f+ [69] Aldehyde dehydrogenase (ALDH1) high [70], or CD44+/CD24- [71] phenotype are also observed in mcf 10A cell line.

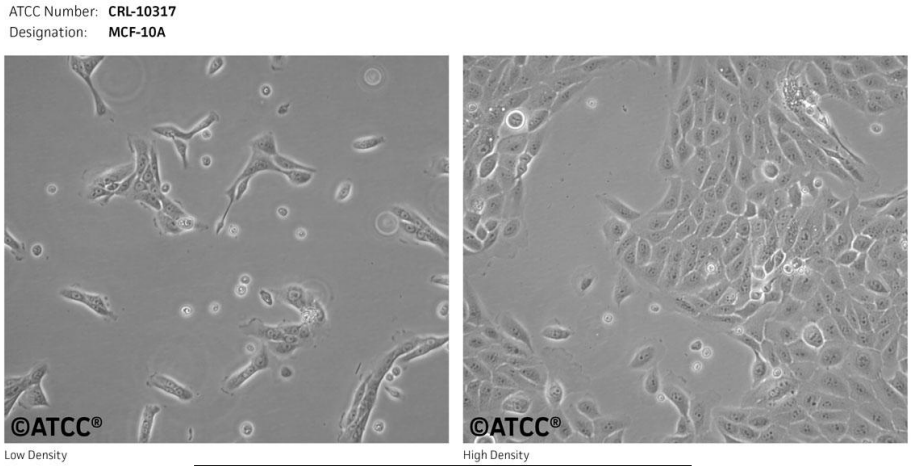

FIGURE 12: MCF 10A In Vitro Cell Culture

3.2 MCF-7

This is a poorly-aggressive and non-invasive cell line [72], normally being considered to have low metastatic potential MCF-7 cells present extensive aneuploidy with important variations in chromosome numbers ranging from 60 to 140 according to the variant examined. Other cytogenetic differences concerned the presence or absence of specific marker chromosomes. The available data suggested an elevated level of genetic instability in MCF-7 cells. The karyotypic differences could reflect changes in selective pressure due to different culture conditions. MCF-7 cells contain a fraction of stem cells able to generate clonal variability. This was proposed as an explanation for the heterogeneity of this cell line and as a model for breast tumor heterogeneity. Different MCF-7 variants undergo divergence at both the genomic and the RNA expression levels.

- CULTURE – Cultured as adherent cell culture in MEM-α (GIBCO) + 10% FBS+ 1% PSL+ 0.1% Insulin. Cell cultured upto 4-5 passages before experimental treatment.
- MORPHOLOGY – Spheroid shaped morphology taken up within 24-36 hours. E-cadherin is the major protein that mediates MCF-7 cell–cell adhesion in spheroids.MCF-7 cells exhibit features of differentiated mammary epithelium: they are positive for epithelial markers, such as E-cadherin, β-catenin and cytokeratin 18 (CK18), and negative for mesenchymal markers, such as vimentin and smooth muscle actin (SMA) [73]. MCF-7 parental cells also maintain the expression of other specific molecular markers of natural epithelial layers, such as claudins and zona occuldens protein 1 (ZO-1), among other proteins that constitute the intercellular junctions [74]. On the other hand, MCF-7 cells are CD44-deficient
- MIGRATION AND INVASION- The parental MCF-7 cells do not usually migrate or invade [72]. An autocrine loop exists for the vascular endothelial growth factor (VEGF) to induce breast cancer cell migration/invasion. MCF-7 cells express lower levels of VEGF than MDA-MB-231 cells, which have high invasive and migration capacities. Without estrogen supplementation, MCF-7 cells do not induce metastasis in mice and have a low capacity of migration in vitro.

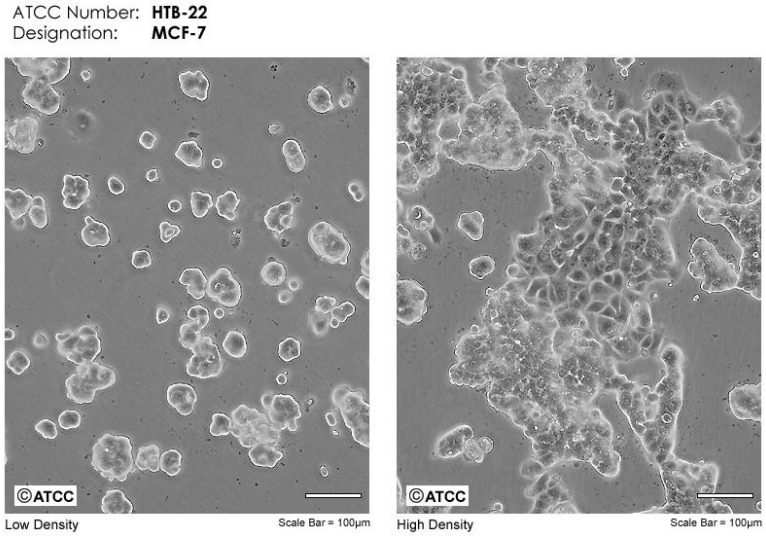

FIGURE 13: MCF 7 IN VITRO CELL CULTURE

3.3 MDA-MB-231

With epithelial-like morphology, the MDA-MB-231 breast cancer cells appear phenotypically as spindle shaped cells. In vitro, the MDA-MB-231 cell line has an invasive phenotype. The MDA-MB-231 cell line is also able to grow on agarose, an indicator of transformation and tumorigenicity, and displays a relatively high colony forming efficiency. In vivo, the MDA-MB-231 cells form mammary fat pad tumors in nude mice. IV injection of cells into the tail vein of nude mice has been shown to produce experimental metastasis.

- CULTURE- MDA-MB-231 culture is done in L-15 media (GIBCO) + 10% FBS+ 1% PSL+ 0.1% Insulin
- MORPHOLOGY- This cell line is ER, PR, and E-cadherin negative and expresses mutated p53. Since the cells also lack the growth factor receptor HER2, they represent a good model of triple-negative breast cancer. In 3D culture, the cell line displays endothelial-like morphology5 and is distinguished by its invasive phenotype, having stellate projections that often bridge multiple cell colonies

- MIGRATION AND INVASION- MDA-MB-231 cells are invasive in vitro and when implanted orthotopically produce xenografts that spontaneously metastasize to lymph nodes. MDA-MB-231 is a highly aggressive, invasive and poorly differentiated triple-negative breast cancer (TNBC) cell line as it lacks oestrogen receptor (ER) and progesterone receptor (PR) expression, as well as HER2 (human epidermal growth factor receptor 2) amplification. As a result of lacking ER and PR expression and HER2 amplification, the cell line was initially classed as a 'basal' breast cancer cell line. However, it is now recognized as belonging to the claudin-low molecular subtype as it exhibits down-regulation of claudin-3 and claudinin-4, low expression of the Ki-67 proliferation marker, enrichment for markers associated with the epithelial-mesenchymal transition and the expression of features associated with mammary cancer stem cells (CSCs), such as the CD44+CD24-/low phenotype. Sub clones of MDA-MB-231 cells that preferentially metastasize either to the bones, brain and lungs of mice following intraventricular injection have also been isolated, thus allowing this cell line to be used in the identification of genes and pathways that are potential mediators of metastasis to specific sites.

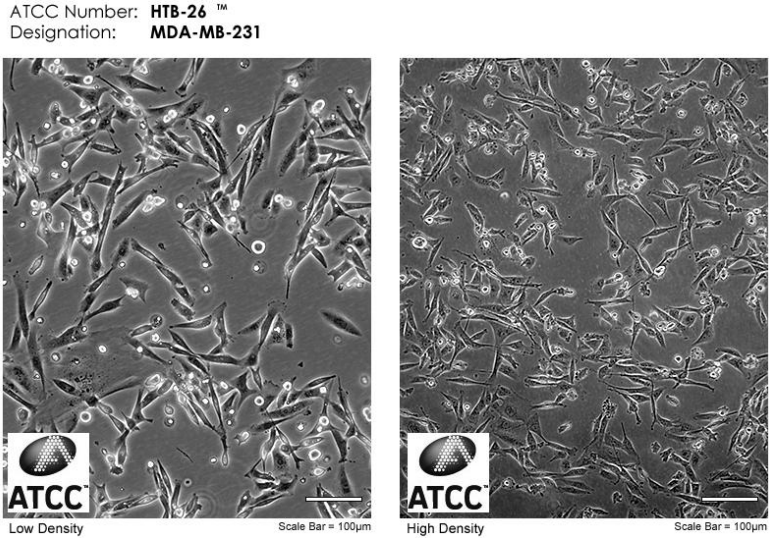

FIGURE 14: MDA-MB-231 In Vitro Cell Culture

OBJECTIVE

To find the role of novel lncRNA PCAT19 in tumorigenesis pathway and its association with Non-muscle myosin II in human breast cancer model.

METHODOLOGY

1. Cell culture and maintenance

2. RNA isolation (by Trizol method)

3. Qualitative and quantitative assay of RNA

4. DNase treatment and cDNA synthesis

5. Qualitative assay of primers of target genes

6. PCR

7. Agarose gel electrophoresis

8. Gene silencing by employing siRNA treatment of target genes

1. CELL CULTURE AND MAINTENANCE

INTRODUCTION:

MCF- 10A, MCF-7 and MDA-MB-231 cell lines were cultured for the experiments by following adherent cell culture protocol. The media requirements for the aforementioned cell lines are tabulated below:

CELL LINE	MEDIA	ADDITIONAL ESSENTIAL FACTORS	STORAGE	INCUBATION CONDITIONS
MCF-10A	D-MEM-F12:HAM in 1:1 ratio	100 ng/ml cholera toxin + 20 ng/ml epidermal growth factor (EGF) + 0.01 mg/ml Human Recombinant Insulin + 500 ng/ml hydrocortisone + 5% gamma irradiated horse serum	4°C	37 ° C + 5% CO_2
MCF- 7	MEM-α	10% FBS+ 1% PSL+ 0.01 mg/ml Human Recombinant Insulin	4°C	37° C+ 5% CO_2
MDA-MB-231	L-15	10% FBS+ 1% PSL+ 0.01 mg/ml Human Recombinant Insulin	4°C	37°C

CELL SEEDING:

- Respective media for the cell line was taken and incubated at 37°C water bath for few minutes for warming up.
- One large culture plate (60 cm^2) and one small culture plate (20 cm^2) were taken under a UV treated sterile animal cell culture hood.
- 7 ml media was poured on the 60 cm^2 culture plate and 3 ml on the 20 cm^2 culture plates.
- Cryovials of the respective cell line was taken out from liquid nitrogen (-196 °C) and immediately placed in 37°C water bath for a few minutes until it was completely thawed.
- For DMSO removal, the cells were centrifuged and the pellet resuspended in fresh media.
- A 5 ml pipette was used to distribute the cells in 2:1 ratio to the 60 cm^2 and 20 cm^2 culture plates respectively.
- The culture plates were gently stirred in clockwise and anti-clockwise manner to distribute cells evenly.
- The culture plates were observed under compound microscope to check for the even distribution.
- Culture plates were incubated at 37° C with 5% CO_2 environment (Not for MDA-MB-231 cell line) for the next 24 hours for proper seeding and adherence.
- Media change was done after 24 hrs by sucking the old media and then adding the same amount of fresh media as before into the respective culture plates.
- Cell-confluency over 80-90% was seen after 72-96 hours incubation.
- Monitoring of cell morphology and media change was done at regular intervals.

SUB-CULTURING:

Cell sub-culturing is done by cell splitting method to increase the capacity, vigor and health of a cell culture. The key component for cell splitting is Trypsin (0.05%). Once trypsin is added to a monolayer of adherent cell culture, the enzyme digests the extra-cellular matrix that binds all the cells of the culture into a single layer. As a result single cells and small clumps of cells float up to the media. The clumps are resolved later by clockwise and anticlockwise movement and hitting of the plates into single cells. These single cells are pelleted down and then seeded into fresh culture plates.

PROCEDURE:

- When cells attain a confluency of 90-100%, it is understood that the cells are ready for splitting
- For, 60 cm^2 culture plate, it was brought to the culture.
- The media was sucked and Trypsin-EDTA (0.05%) was pipette onto both the culture plates
- The plates were placed in incubator for 2 mins.
- The plates were swirled clockwise and anti-clockwise and hit for quite a few times to uplift trypsinized cells.
- Whether cells were uplifted, was checked under the microscope.
- The trypsin was deactivated by addition of fresh media (2 ml for 60 cm^2 plate and 1 ml for 20 cm^2 plate) onto each of the culture plates.
- The cell suspension was then collected into a sterile round bottom tube (RBT) and centrifuged at 3000 rpm for 2 mins
- The cells were pelleted down at the bottom of the RBT.
- The supernatant was discarded and the pellet was resuspended in fresh media.
- The suspension was stiturated for breaking down the pellet into single cells.
- Few drops of the suspension were added into a sterile culture plate containing fresh media
- To the rest 5% DMSO was added and mixed well.
- The DMSO added cell suspension was distributed into properly labeled sterile cryovials for further preservation if required.
- Cryovials were taken into 1º cooler and placed into -80º C freezer for 24 hours.
- The cryovials were finally stored in liquid nitrogen until further use.

2. RNA ISOLATION (BY TRIZOL METHOD)

PRINCIPLE:

RNA (Ribonucleic acid) is mostly single stranded molecules composed of phosphate and ribose sugar, nitrogen bases adenine, guanine, uracil, cytosine which are bound with ribose sugars. Trizol Reagent (Guanidinium thiocyanate and acid phenol) is a reagent used for isolating RNA from cells and tissues. Cells are washed (with PBS) and then lysed by adding Trizol. It works by maintaining the integrity of RNA molecules during tissue homogenization, while disrupting and breaking down cellular components at the same time. (Trizol inactivates RNAses.) Then, further lysis is aided by pipetting or centrifugation (for tissues). Chloroform separates the sample solution into an organic phase and an (RNA-containing) aqueous phase, which is collected. The phases after Chloroform addition are:

i. Bottom layer: Contains proteins, lipids and cell debris (phenol Chloroform phase)
ii. Middle layer: Contains DNA
iii. Top layer: Colorless aqueous phase, contains RNA

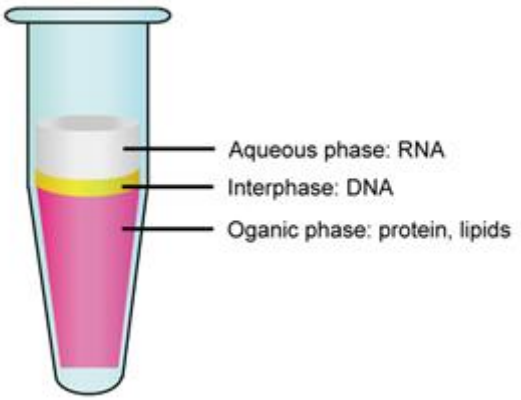

FIGURE 15: Separation Of Aqueous Phase And Organic Phase In Total RNA Extraction Method By Trizol Method

MATERIALS:

- DEPC-Treated water.
- Trizol reagent
- Ice cold PBS
- 70% ethanol
- Isopropyl alcohol
- Chloroform(taken in a falcon tube for phase separation)

PROCEDURE:

- 90-100% confluent cell plates of $20cm^2$ were taken. Cell scrapping was done by a sterile scrapper for the homogenized cell lysate preparation.
- 0.2ml of chloroform was added per 1ml of Trizol reagent.
- The samples were vortexed vigorously for 15seconds and then incubated at room temperature for 2-3 minutes.
- The samples were centrifuged at 4°C and 12000g for 15 minutes.
- The sample got separated into 3 distinct phases where the upper aqueous layer was collected for RNA without intervening intermediate and the bottom phase during collection.
- 0.5ml of isopropyl alcohol was added and incubated for 10 minutes at room temperature.
- The samples were centrifuged at 4°C and 12000g for 10 minutes.
- The supernatant was decanted and 75% ethanol was added to wash the RNA precipitate.
- The samples were centrifuged at 4°C and 7500g for 5 minutes. This washing step was performed twice.
- Then the precipitate was air dried under laminar air-flow and dissolved in 20µl of molecular biology grade water(MBGW)

3. QUALITATIVE AND QUANTITATIVE ASSAY OF RNA

- **QUANTITATIVE ASSAY OF RNA BY SPECTROPHOTOMETRIC ANALYSIS**

  RNA sample absorbance is determined on the spectrophotometer at 260nm. 260nm is the highest absorbance wavelength for all nucleic acid (DNA, RNA, nucleotides). If the ratio of absorbance at 260:280 is 2.0, the RNA isolated is considered to be pure.

- **QUANTITATIVE ESTIMATION OF RNA**

  Since nucleic acids absorb UV light maximally at 260nm, for RNA OD_{260} is taken.

  Conversion factor = 40

  The concentration in the sample is calculated by using the formula:

  $$A_{260} \times \text{dilution} \times 40 = [\text{RNA}]\ \mu g/ml$$

- **QUALITATIVE ANALYSIS OF RNA- (RNA GEL)**

  The overall quality of RNA isolation was assessed by electrophoresis on 1% agarose gel. The analysis gives information about RNA yield and whether it contains DNA contamination. Intense care is needed for the qualitative analysis of RNA to avoid RNA contamination.

MATERIALS**:**

- DEPC treated water
- TBE buffer in DEPC treated water
- Agarose
- EtBr
- Mixed mass DNA ladder (80-10,000) bp

PROCEDURE:

- TBE buffer was prepared with DEPC water.
- 1% agarose gel was prepared (0.45 g agarose powder + 2 µl EtBr) in 45 ml of TBE buffer.
- Gel casting was done while the agarose is warm and comb was inserted before pouring agarose onto gel-cast.
- The apparatus is left to a standstill for about an hour till the gel solidifies.
- 8 µl of RNA sample (conc- 1 µg/ µl) is mixed with 2 µl of 5x DNA loading dye. 10 µl sample is taken for loading onto the wells of the solidified agarose gel.
- 4 µl of mixed mass DNA ladder is added to one of the lanes
- The gel is run for 1 hr 15 minutes on 85 V
- After gel run is over, the gel image was captured using Chemi-doc system under UV light.
- Image was analyzed for the next plan of work.

4. DNase TREATMENT AND cDNA SYNTHESIS

In genetics, **complementary DNA** (**cDNA**) is DNA synthesized from a single-stranded RNA template in a reaction catalyzed by the enzyme reverse transcriptase. This enzyme, which naturally occurs in retroviruses, operates on a single strand of mRNA, generating its complementary DNA based on the pairing of RNA base pairs (A, U, G and C) to their DNA complements (T, A, C and G, respectively) using short random primers. The synthesized cDNA can be used in gene cloning, as gene probes, in the preparation of cDNA library and also as a template for amplification by PCR. Since we convert the RNA back into DNA via reverse transcriptase, and then attempt to amplify target genes, any genomic DNA present could also be amplified, leading us to question whether the final results of RT-PCR are due to the cDNA that we generated, or simply the genomic DNA contaminating our sample. To guard against this, we treat RNA samples with DNase I, an enzyme that selectively degrades DNA.

4.1 DNase TREATMENT

MATERIALS:

- MBGW
- DNaseI
- 10X Reaction buffer
- STOP solution (25mM EDTA)
- Thermal cycler

PROCEDURE:

- 1 µg/µl sample was taken in 7 µl MBGW
- 1 µl of 10x reaction buffer (RT buffer) was mixed with 1 µl DNase I
- The tubes were mixed well and spun at 3000 rpm for 2 mins
- The tubes were incubated at room temperature for 15 mins
- 1 µl of STOP solution was added
- The tubes were given a free spin and then incubated at 70°C for 8 mins in Thermal cycler and chilled on ice.

4.2 cDNA SYNTHESIS

MATERIALS AND METHODS:

- Preparation Of cDNA Mastermix

➢ For two samples, the following reagents in their respective amounts were calculated and added accordingly:

REAGENTS	AMOUNT (µl)
10X RT BUFFER	2
25X DNTP MIX (100MM)	0.8
10X RT RANDOM PRIMERS	2
RNASE INHIBITOR	1
NUCLEASE-FREE H_2O	3.2
TOTAL	10

➢ 9 µl of the solution is aliquot was a centrifuge tube and labeled as negative control.

➢ 1 µl of nuclease free water was added to it instead of reverse transcriptase.

➢ 1 µl of reverse transcriptase for each sample was added to the main mastermix.

➢ The samples were mixed well and spun.

➢ 10µl of DNase treated RNA was added to each tube. So, the final volume in each tube became 20 µl.

➢ The tubes were placed in thermal cycler for the reverse transcription reaction to take place.

- Reverse Transcription In Thermal Cycler

STEP	TEMPERATURE	TIME
STEP 1	25°C	10 MINS
STEP 2	37°C	120 MINS
STEP 3	85°C	5 MINS
FREE SPIN		
STEP 4	-20°C	FOREVER

5. QUALITATIVE ASSAY OF PRIMERS OF TARGET GENES

MATERIALS:

- 15% PAGE
- 1x TBE
- 10x TBE
- Primer Main Stock (Forward and Reverse)

PROCEDURE: PREPARATION OF BUFFER

- 1X TBE (FOR 1 L)

REAGENTS	AMOUNT
TRIS	10.8 g
BORIC ACID	5.5 g
0.5 M EDTA	4 ml
H_2O	1 l
ETBR	50 µl

- 10X TBE (50 ml)

REAGENTS	AMOUNT
TRIS	5.4g
BORIC ACID	2.75g
0.5 M EDTA	2 ml
H_2O	50 ml

- PREPARATION OF 15% PAGE

REAGENT	AMOUNT
30% POLYACRYLAMIDE	5 ml
10X TBE	1 ml
10%APS	40 µl
TEMED	6 µl
DEPC H2O	4 µl

- The main stocks of forward and reverse primer were taken.
- 16 µl from each stock was mixed with 4µl of 5x DNA loading dye.
- The sample was loaded onto agarose gel and electrophoresed at 95V for about 1 hour.

6. POLYMERASE CHAIN REACTION (PCR)

Polymerase chain reaction (PCR) is a method widely used in molecular biology to make many copies of a specific DNA segment. Using PCR, a single copy (or more) of a DNA sequence is exponentially amplified to generate thousands to millions of more copies of that particular DNA segment. The vast majority of PCR methods rely on thermal cycling.[1] Thermal cycling exposes reactants to repeated cycles of heating and cooling to permit different temperature-dependent reactions—specifically, DNA melting and DNA polymerase enzyme-driven DNA replication. PCR employs two main reagents - primers (which are short single strand DNA fragments known as oligonucleotides that are a complementary sequence to the target DNA region) and a DNA polymerase. In the first step of PCR, the two strands of the DNA double helix are physically separated at a high temperature in a process called DNA melting. In the second step, the temperature is lowered and the primers bind to the complementary sequences of DNA. The two DNA strands then become templates for DNA polymerase to enzymatically assemble a new DNA strand from free nucleotides, the building blocks of DNA. As PCR progresses, the DNA generated is itself used as a template for replication, setting in motion a chain reaction in which the original DNA template is exponentially amplified.

MATERIALS:

- Thermal cycler
- cDNA samples
- Mastermix
- Forward and Reverse primers
- dH_2O

PROCEDURE:

- PCR PROTOCOL

SAMPLE	MASTERMIX (µl)	PRIMER FORWARD (µl)	PRIMER REVERSE (µl)	DH_2O (µl)	CDNA (µl)
S1	12.5	1	1	7.5	3

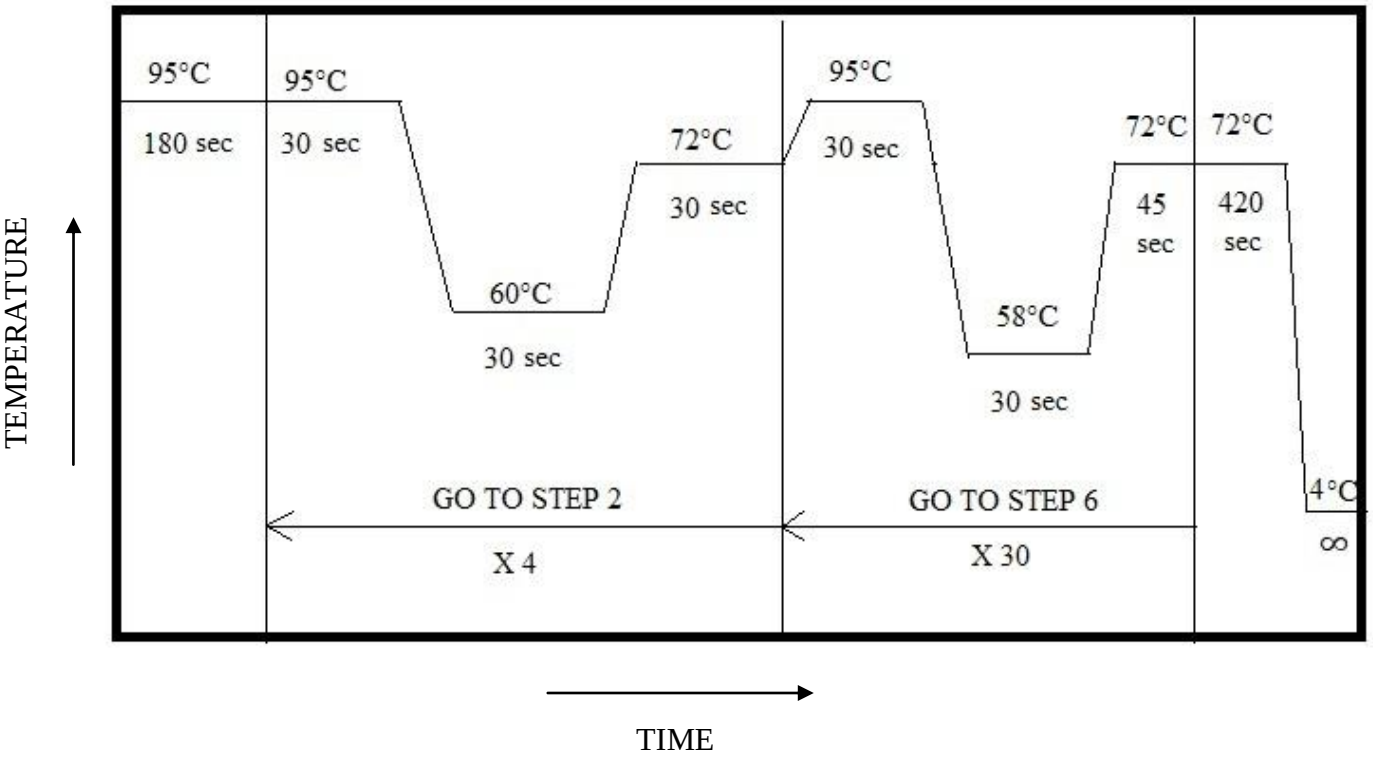

FIGURE 16: Sample PCR Programme (Human β-Actin)

- The samples were placed in the thermal cycler and programme set. Once the programme ended, the samples were retrieved and stored at 4°C till further use.

7. AGAROSE GEL ELECTROPHORESIS

Agarose gel electrophoresis is a technique used to separate nucleic acids that differ in size. When charged molecules are placed in an electric field, they migrate towards the opposite pole based on their charge. For example nucleic acid has a consistent negative charge imparted by their phosphate backbone, and migrates toward the anode. Gel electrophoresis is usually performed for analytical purposes, often after amplification of DNA via PCR. Agarose is a polysaccharide extracted from seaweed. It is typically used at concentrations of 0.5 to 2%. Higher agarose concentration is inversely proportional to the pore size. According to the size of DNA, the concentration of agarose is adjusted following the standard table.

MATERIALS

- Electrophoresis buffer (1X TAE)
- Agarose
- DNA ladder
- Gel caster
- Ethidium bromide solution
- 5x DNA Loading dye
- Gel combs
- DNA gel electrophoresis system

PROCEDURE

- PREPARATION OF 1X TAE BUFFER (1 L)

TRIS BASE	4.84 gm
GLACIAL ACETIC ACID	1.14 ml
EDTA	2 mM

- 2% agarose gel was prepared by taking 1gm of agarose powder and mixing it 50ml of TAE Buffer.
- Then it was heated for around 1minute and allowed to cool.
- After that, 2µl Ethidium Bromide was added to it.
- Then molten gel was cast on the electrophoresis tray and allowed to solidify.
- The PCR products were taken. PCR product = 25 µl and 5 µl of 5x DNA loading dye added to it to make the total volume as 30 µl. 20 µl from there was loaded onto the agarose gel wells and electrophoresed at 90 V for less than 1 hr (45-50 minutes).

8. GENE SILENCING BY EMPLOYING siRNA TREATMENT OF TARGET GENES

Small interfering RNA (siRNA), known as short interfering RNA or silencing RNA, is a class of double-stranded RNA molecules, 20-25 base pairs in length and operating within the RNA interference (RNAi) pathway. It interferes with the expression of specific genes with complementary nucleotide sequences by degrading mRNA after transcription, preventing translation [35]. Although is an intracellular product, artificially synthesized custom made siRNA are also used widely for genetic modification in cells. siRNA transfection is a powerful tool used to understand the mechanisms of gene regulation and molecular pathways.

Since MCF 7 cell line showed more expression of the target gene PCAT19, so MCF 7 cell line was chosen for study by gene silencing. Here, silencing was introduced for the target genes, PCAT19 to check the variation in the level of its expression in MCF 7 cells. Two different siRNA for PCAT19 was introduced in MCF 7 cells to check for the expression of PCAT19 under siRNA treatment and control.

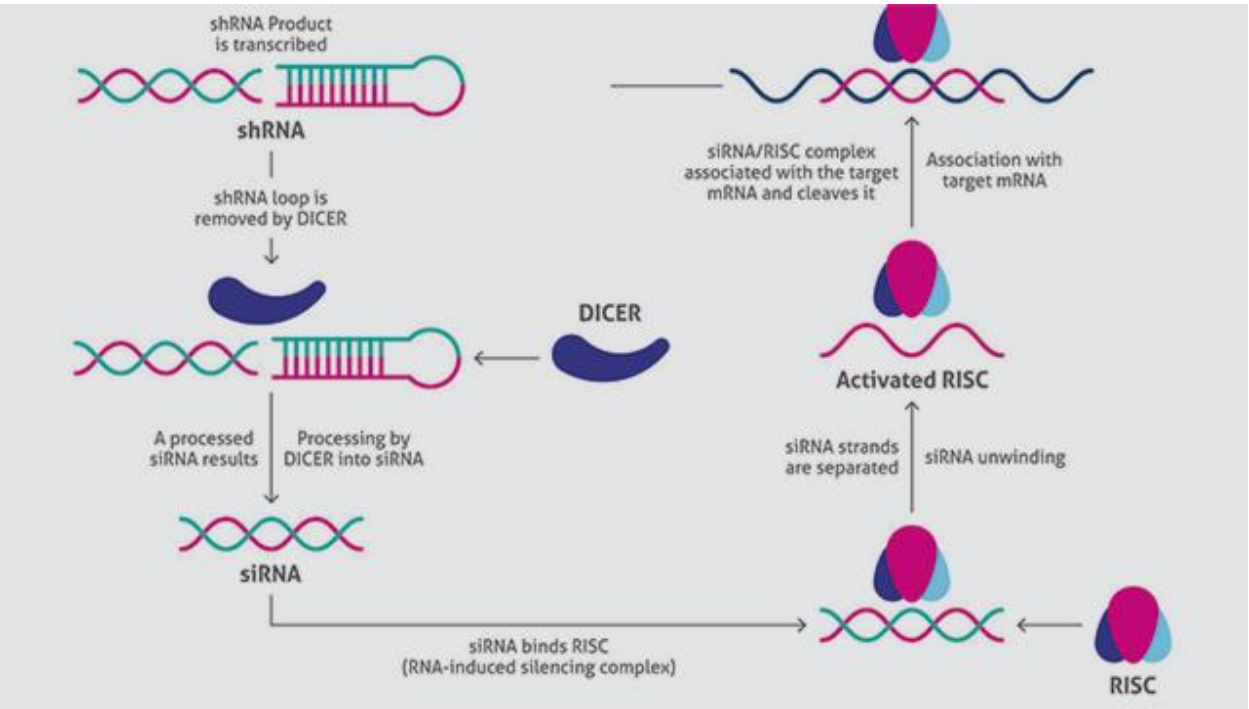

FIGURE 17: Mechanism Of Function Of siRNA

MATERIALS AND REAGENTS REQUIRED

- Lipofectamine RNAimax Reagent (Invitrogen)
- Opti-MEM media (Invitrogen)
- Customised synthetically prepared siRNA against proposed lncRNAs- HOTAIR and PCAT19 fluorescent labelled with MISS SIRNA FLUOR UNIV NEG CNTRL #1 CYAN I (Sigma Aldrich)
- Cell culture of MCF-7 and MDA-MB-231 at 60-70% confluency.

For 20 cm^2 culture plate:

REAGENT	AMOUNT
OPTIMEM MEDIA	150µL
LIPOFECTAMINE RNAIMAX	9 µL
OPTIMEM MEDIA	150 µL
SIRNA	3 µL
MIXTURE OF LIPOFECTAMINE RNAiMAX + SIRNA (OPTIMEM MEDIA) IN 1:1 RATIO	

PROCEDURE

- 20 cm^2 cell plate taken with 60-70% confluency
- Lipofectamine RNAiMAX reagent is diluted in OptiMEM media
- The siRNA mother stock is first prepared and from there a working stock is prepared for transfection in OptiMEM media such that transfected siRNA will have a concentration of 10µM/µl.
- Diluted siRNA and diluted lipofectamine RNAiMAX reagent are mixed in a 1:1 ratio
- The complete mixture of reagents is incubated at room temperature for 5 minutes.
- Media is sucked from the plate and siRNA- lipid complex was added to cell plates
- The cell plates were incubated for 6 hours in 37 ° C.
- Media change was done and again kept in incubator for the next 24 hours.

RESULTS

1. RNA GEL ELECTROPHORESIS

RNA gel electrophoresis was done showing three clear bands of 28S, 18S and 5S for the three cell lines: MCF 10A, MCF 7, MDA-MB-231.

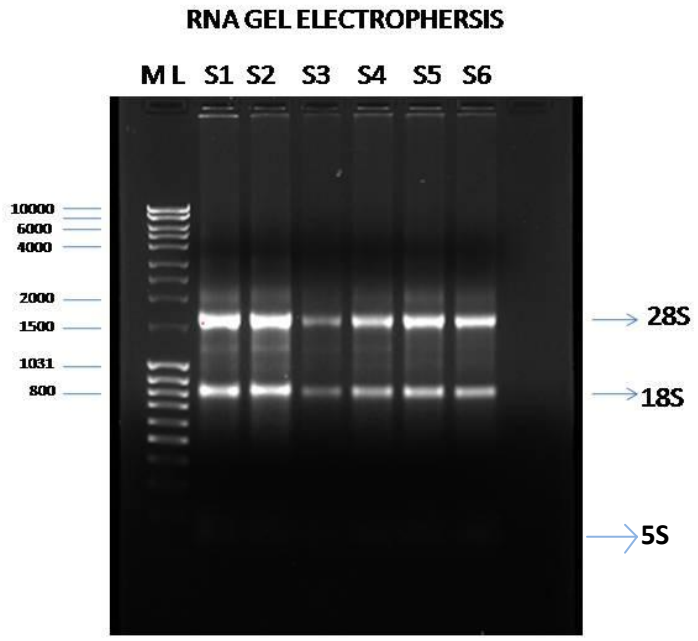

Figure 18: RNA Agarose Gel Electrophoresis of Breast Cancer Cell Lines S1=MCF 10A, S2=MCF 10A, S3=MCF 7, S4=MCF7, S5= MDA-MB231, S6= MDA MB 231

2. REVERSE TRANSCRIPTION PCR (RT-PCR)

Reverse transcription polymerase chain reaction for PCAT19, HOTAIR and NM IIC was done.

- Expression of MYH14 was found to be approximately 7x fold higher in MCF 7 cell line (lane 2 and 3 of Fig. 19) as compared to the metastatic cell line MDA-MB-231(lane 1 in fig. 19). No bands were observed in MCF 10A (lab data)
- Expression of MYH14 was found to be approximately 6x fold higher in MCF 7 cell line (lane 2 and 3 of Fig. 20) as compared to the metastatic cell line MDA-MB-231(lane 1 in fig. 20). No bands were observed in MCF 10A. (lab data)
- Expression of MYH14 was found to be approximately 16x fold higher in MCF 7 cell line (lane 2 and 3 of Fig. 21) as compared to the metastatic cell line MDA-MB-231(lane 1 in fig. 21). No bands were observed in MCF 10A. (lab data)

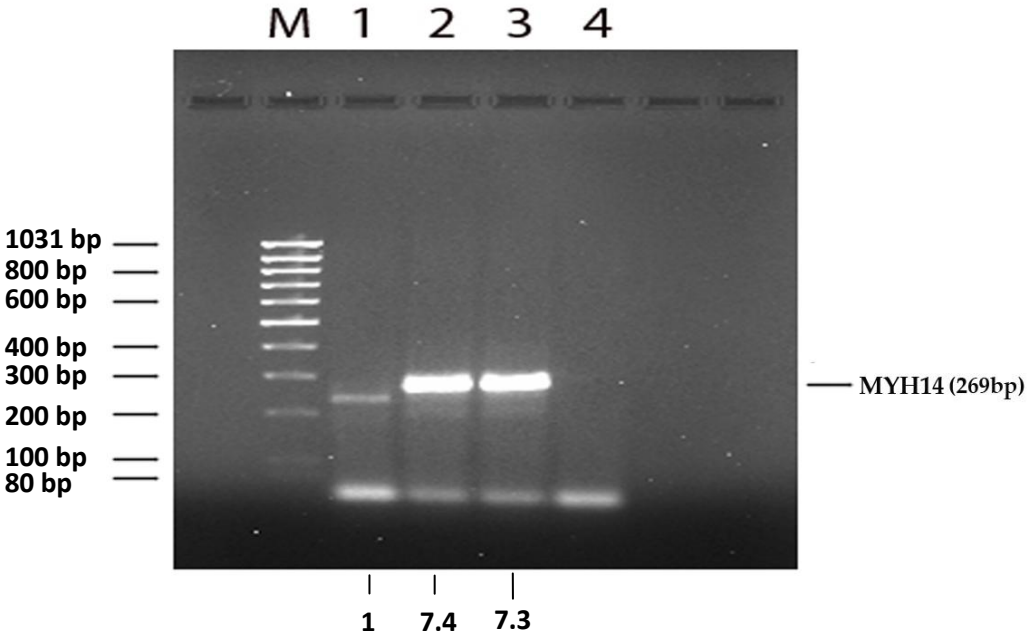

FIGURE 19: Detection of MYH14 Gene By RT-PCR. PCR Product Size= (269 bp) On EtBr Stained Agarose Gel. M= Mass Ruler DNA Ladder Low Range (80-1000bp), 1= MDA-MB-231, 2 AND 3= MCF 7, 4= Negative Control

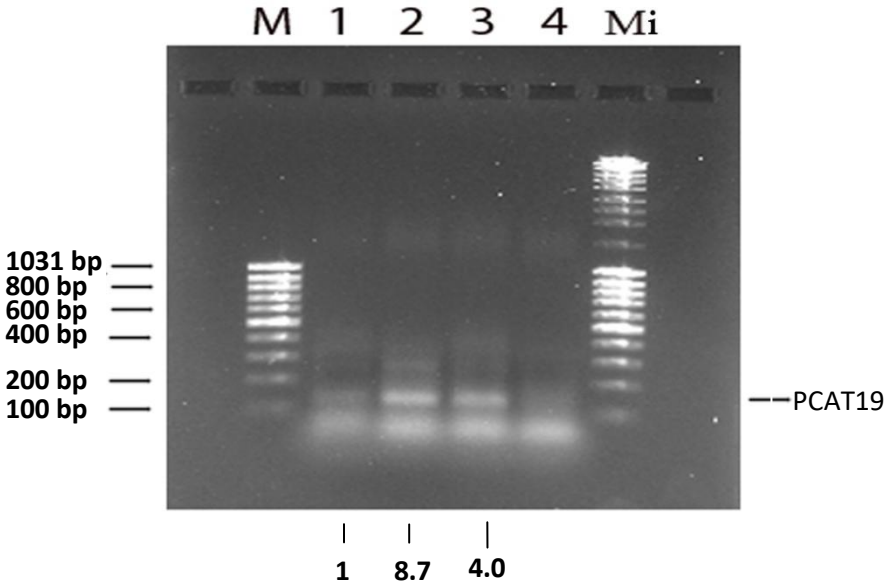

FIGURE 20: Detection of PCAT19 Gene By RT-PCR.PCR Product Size= (126 bp) On EtBr Stained Agarose Gel. M= Mass Ruler DNA Ladder Low Range (80-1000bp), Mi= Mass Ruler DNA Ladder Mix (10 Kb) 1= MDA-MB-231, 2 AND 3= MCF 7, 4= Negative Control

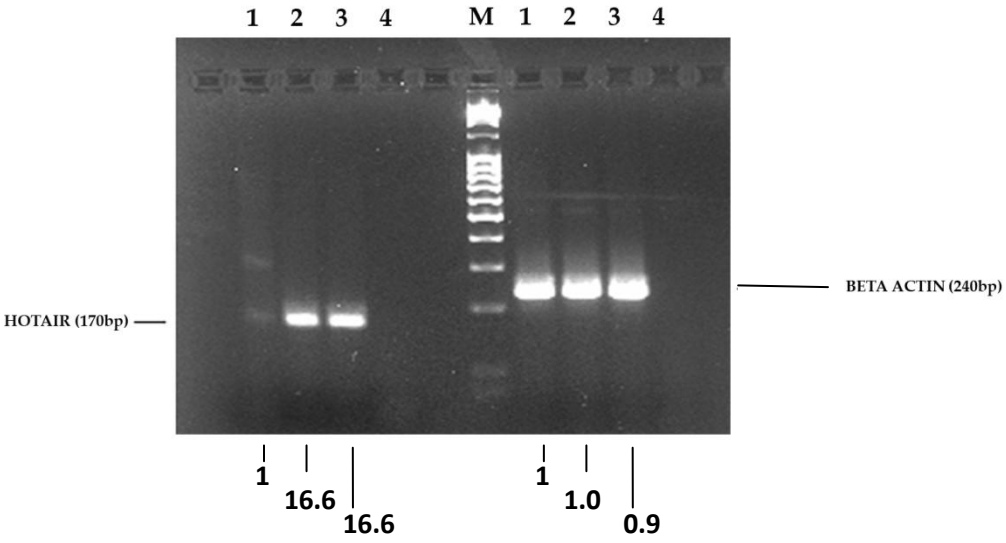

FIGURE 21: Detection of HOTAIR and Human β Actin Gene By RT-PCR. PCR Product Size= (240 bp) On EtBr STAINED AGAROSE GEL. M= Mass Ruler DNA Ladder Low Range (80-1000bp), 1= MDA-MB-231, 2 AND 3= MCF 7, 4= Negative Control

- GENE SILENCING VIA siRNA TREATMENT

    - Gene silencing via siRNA treatment was done for the gene PCAT19 with a transfection efficiency of 95% using Lipofectamine RNAimax Reagent. Treatment showed reduced expression of PCAT19 gene found from RT-PCR analysis.*
    - The RT-PCR analysis also showed reduced expression of NM IIC. Both siRNAs specific for PCAT19 treated samples responded for the gene NM IIC in a positive correlation.*

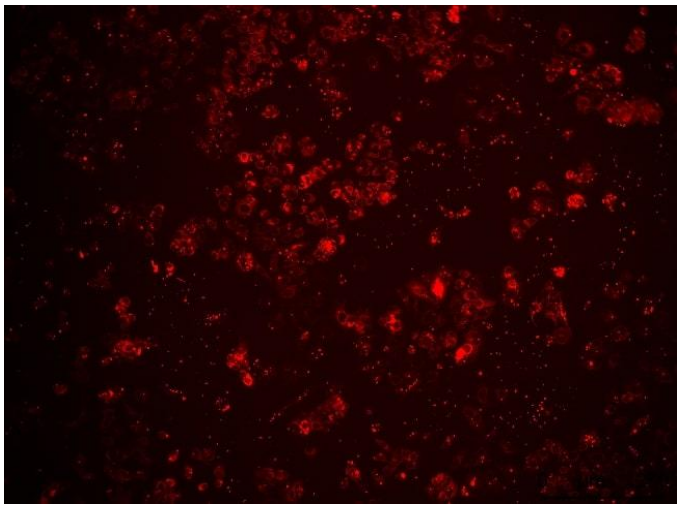

FIGURE 22: Non-specific siRNA treated MCF7 cells under Green light of Fluorescence Microscope at 10x .magnification. (Negative control to check transfection efficiency)

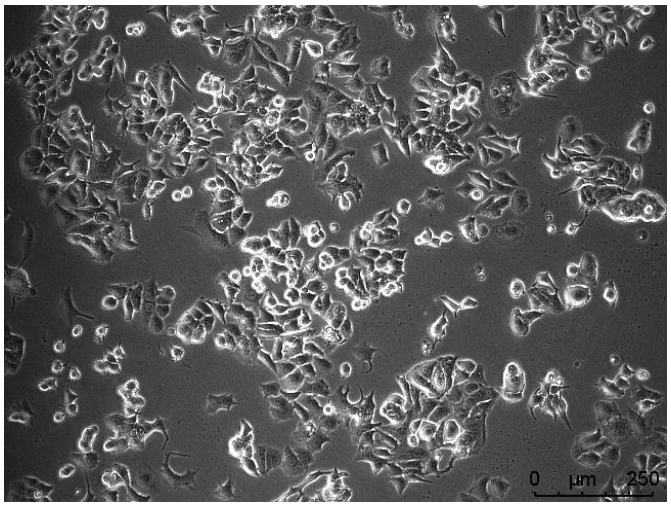

FIGURE 23: Non-specific siRNA treated MCF7 cells under gray scale of Fluorescence Microscope at 10x magnification. (Negative control to check transfection efficiency)

*Data not released since it is under publication.

DISCUSSION

In the human genome, the ratio of non-coding DNA to total genomic DNA is approximately 98.5%. Recent studies have revealed that transcription is not restricted to protein-coding regions, but occurs throughout the genome (>90%), including non-coding regions [79]. This yields large numbers of non-coding RNAs. One such group of non- coding RNAs is the lncRNAs. Although our current understanding of the functional role of lncRNAs is limited, several lncRNAs have been identified for being associated with the progression of cancer, namely metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), HOX antisense intergenic RNA (HOTAIR), antisense non-coding RNA in the INK4 locus (ANRIL) [76] and PCAT19 overexpressed in prostate cancer. PCAT19 lncRNA was also found to be overexpressed in tumorigenic laryngeal tissues [48] and tumorigenic gall bladder tissues [50] as compared to their normal and metastatic tissues respectively. However no report of its expression in breast cancer has been found yet. The motor protein non-muscle myosin II (NMII) acts as a master regulator of cell morphology, with a role in several essential cellular processes, including cell migration and post-synaptic dendritic spine plasticity in neurons. NMII also generates forces that alter biochemical signaling, by driving changes in interactions between actin-associated proteins that can ultimately regulate gene transcription. In addition to its roles in normal cellular physiology, NMII has recently emerged as a critical regulator of diverse, genetically complex diseases, mainly cancer.[78] In mammals, three different isoforms of nonmuscle myosin II : II-A, II-B and II-C, are found. These isoforms generate due to the difference in the length of the heavy chain of the myosin protein. While a few cells contain a single isoform, most contain more than one, including isoforms generated by alternative splicing. In humans, these isoforms are encoded by three different genes, myh9 (NM IIA), myh10 (NM II-B) and MYH14 (NM II-C), present on three different chromosomes [11, 12]. It was found that the PCAT19 lncRNA coding gene resides in cis position with the MYH14 gene coding for NM IIC [77]. Therefore, we hypothesized that PCAT 19 lncRNA may have a role in regulating NM IIC in cis position in breast cancer model. We proceeded the hypothesis by dealing with three human breast cell lines featuring normal, tumorigenic and metastatic phenotype MCF 10A, MCF 7 and MDA-MB-231 respectively. Quantitative gene expression analysis was done with three above mentioned cell lines for HOTAIR, PCAT19 and NM IIC. The result of reverse transcription PCR showed the higher expression level of the HOTAIR, PCAT19 and NM IIC in MCF 7 (tumorigenic) cell line than MDA-MB-231(metastatic). The RT-PCR showed no expression for the target genes in MCF 10A.This enables us to conclude that expression of PCAT19 and NM IIC might have a role in tumorigenesis and once metastasis is established, its expression is lowered down. To verify and validate our results further, we selected the MCF 7 cell line and performed gene silencing of the PCAT 19 gene. The results showed reduced expression of PCAT19 in RT-PCR. This siRNA treated sample of MCF 7 also showed reduced expression of NM IIC. Hence, it can be concluded that the genomic study indicates that there is a positive correlation between the PCAT19 lncRNA and NM IIC protein in breast cancer progression and tumorigenesis. However, the functional or regulatory mechanism of PCAT19 lncRNA over NM IIC in breast cancer progression remains to be discovered in near future.

FUTURE PROSPECTS

PCAT19 is a family of long non-coding RNAs involved in cell migration, cell division and tumorigenesis. The effects of inhibition of this lncRNA could be observed by a PCAT19 gene knock-out. Given an opportunity I would like to establish the importance of PCAT19 in better prognosis for breast cancer studies and how PCAT19 regulates MYH14 by ChIP assay. Since, the MYH14 promoter is known from published literature so transcription factor profiling associated with MYH14 expression could be done taking in consideration the lncRNA PCAT19.Also, HOTAIR lncRNA gene could be silenced with the help of similar siRNA treatment and its variation of expression monitored to be correlated with NM IIC and PCAT19 genes. Additionally, after more study, verification and validation, PCAT19 could be established as a potential therapeutic target for breast cancer.

REFERENCES

1. Anca Maria Cîmpean, "Victor Babeş" University of Medicine and Pharmacy, *The Prognostic Role Of Mast Cells According With The Histological Growth Pattern In Liver Metastasis Of Tumors With Digestive Origins,* Timişoara, Timiş, Romania, J Cancer 2017; 8(16):3131-3141. doi:10.7150/jca.18457

2. Dai X, Cheng H, Bai Z, Li J. *Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping.* J Cancer2017; 8(16):3131-3141. doi:10.7150/jca.18457

3. Shuya Yano, Kiyoto Takehara, Shinji Miwa,Hiroyuki Kishimoto, Hiroshi Tazawa, Yasuo Urata, Shunsuke Kagawa, Michael Bouvet, Toshiyoshi Fujiwara and Robert M. H,offman, National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi, China; 2. The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China, *Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping*, J Cancer 2017; 8(16):3131-3141. doi:10.7150/jca.18457

4. Mary Anne Conti, Sachiyo Kawamoto, And Robert S. Adelstein National Heart, Lung, and Blood Institute, National Institutes of Health, 10 Center Dr MSC 1762, Building 10, Room 8N202, Bethesda, MD 20892-1762, USA, *Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping*, J Cancer 2017; 8(16):3131-3141. doi:10.7150/jca.18457

5. Yanyan Tang1,2† , Yi He1,2† , Ping Zhang2,6† , Jinpeng Wang2 , Chunmei Fan2 , Liting Yang2 , Fang Xiong5 , Shanshan Zhang2,3, Zhaojian Gong2 , Shaolin Nie1 , Qianjin Liao1,2, Xiayu Li2,4, Xiaoling Li2,3, Yong Li5 , Guiyuan Li1,2,3, Zhaoyang Zeng1,2,3, Wei Xiong1,2,3* and Can Guo1,2,3* , 1 Department of Colorectal Surgery, Hunan Cancer Hospital and The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, Hunan, China, Tang et al., *LncRNAs regulate the cytoskeleton and related RHO/ROCK signaling in cancer metastasis.* Molecular Cancer (2018) 17:77 https://doi.org/10.1186/s12943-018-0825-x

6. Shouping Xu, Dejia Kong, Qianlin Chen, Yanyan Ping and Da Pang, *Oncogenic long non-coding RNA landscape in Breast Cancer*, Molecular Cancer201716:129 https://doi.org/10.1186/s12943-017-0696-6

7. Wang Y1,2, He H1,2, Li W1,2, Phay J3, Shen R4, Yu L5,6, Hancioglu B6, de la Chapelle A7,2 , PTCSC2 *Binding Protein MYH9 Can Regulate* FOXE1,Proc Natl Acad Sci U S A. 2017 Jan 17*;* 114(3):474-479. doi: 10.1073/pnas.1619917114.

8. Zhang L1, Li S2, Choi YL3, Lee J3, Gong Z2, Liu X4, Pei Y4, Jiang A2, Ye M2, Mao M5, Zhang X6, Kim J7, Chen R8,, MSD China R&D, Beijing, China, *Systematic identification of cancer-related long noncoding*

RNAs and aberrant alternative splicing of quintuple-negative lung adenocarcinoma through RNA-Seq, Lung Cancer. 2017 Jul; 109:21-27. doi: 10.1016/j.lungcan.2017.04.009.

9. Pollard, Thomas D.; Korn, Edward D. (1973). "*Acanthamoeba myosin. I. Isolation from Acanthamoeba castellanii of an enzyme similar to muscle myosin*". The Journal of Biological Chemistry. 248 (13): 4682–90. PMID 4268863. Archived from the original on 2016-01-06.

10. Paola Llinas,1,2,* Tatiana Isabet,1,2,* Lin Song,3,* Virginie Ropars,1,2 Bin Zong,3 Hannah Benisty,1,2Serena Sirigu,1,2 Carl Morris,3 Carlos Kikuti,1,2 Dan Safer,3 H. Lee Sweeney,3,† and Anne Houdusse1,2,† , Anne Houdusse, Institut Curie CNRS, UMR144, 26 rue d'Ulm, 75248 Paris cedex 05, France, *How actin initiates the motor activity of myosin.* Dev Cell. 2015 May 26; 33(4): 401–412.

11. Aguilar, H. N.; Mitchell, S.; Knoll, A. H.; Yuan, X. (2010). *"Physiological pathways and molecular mechanisms regulating uterine contractility"*. Human Reproduction Update. 16 (6): 725–744. doi:10.1093/humupd/dmq016. PMID 20551073

12. Matsuoka, R.; Yoshida, M. C.; Furutani, Y.; Imamura, S. I.; Kanda, N.; Yanagisawa, M.; Masaki, T.; Takao, A. (1993). *"Human smooth muscle myosin heavy chain gene mapped to chromosomal region 16q12"*. American Journal of Medical Genetics. 46 (1): 61–67. doi: 10.1002/ajmg.1320460110. PMID 7684189

13. von der Ecken, J.; Heissler, Sarah M.; Pathan-Chhatbar, S.; Manstein, D.J.; Raunser, S. (2016). "*Cryo-EM structure of a human cytoplasmic actomyosin complex at near-atomic resolution"*. Nature. 534 (7609): 724–728. doi:10.1038/nature18295. PMID 27324845

14. Tyska, Matthew J.; Warshaw, David M. (2002). *"The myosin power stroke"*. *Cell Motility and the Cytoskeleton.* 51 (1): 1–15. doi:10.1002/cm.10014. PMID 11810692

15. Uyeda TQ, Iwadate Y, Umeki N, Nagasaki A, & Yumura S (2011) *Stretching actin filaments within cells enhances their affinity for the myosin II motor domain.* PLoS ONE 6(10):e26200

16. Ebrahim S1, Fujita T, Millis BA, Kozin E, Ma X, Kawamoto S, Baird MA, Davidson M, Yonemura S, Hisa Y, Conti MA, Adelstein RS, Sakaguchi H, Kachar B., Laboratory of Cell Structure and Dynamics, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, MD 20892, USA,*NM II forms a contractile sarcomeric network to regulate apical cell junctions and tissue geometry.* Curr Biol. 2013 Apr 22; 23(8):731-6. doi: 10.1016/j.cub.2013.03.039

17. Aibing wang, Xuefei Ma, Mary Anne Conti and Robert S. Adelstein (2011). *Distinct and redundant roles of the non-muscle myosin II isoform and functional domains.* Biochem. Soc. Trans. 39, 1131-1135.

18. Vicente-Manzanares, M., Ma, X., Adelstein, R. S., and Horwitz, A. R. (2009) *Non-muscle myosin II takes centre stage in cell adhesion and migration.* Nat. Rev. Mol. Cell Biol. 10, 778-790.

19. Milton, D. L., Schneck, A. N., Ziech, D. A., Ba, M., Facemyer, K. C., Halayko, A. J., Baker, J. E., Gerthoffer, W. T., and Cremo, C. R. (2011) *Direct evidence for functional smooth muscle myosin II in the 10S self-inhibited monomeric conformation in airway smooth muscle cells.* Proc. Natl. Acad. Sci. U. S. A. 108, 1421-1426
20. Yanyan Tang†, Yi He†,Ping Zhang†, Jinpeng Wang, Chunmei Fan, Liting Yang, Fang Xiong, Shanshan Zhang, Zhaojian Gong, Shaolin Nie, Qianjin Liao, Xiayu Li, Xiaoling Li,YongLi, Guiyuan Li,Zhaoyang Zeng,Wei Xiong and Can Guo, *LncRNAs regulate the cytoskeleton and related RHO/ROCK signaling in cancer metastasis.* Molecular Cancer (2018) Molecular Cancer201817:77 https://doi.org/10.1186/s12943-018-0825-x

21. Trybus, K, M., and Lowey, S. (1984*) Conformational states of smooth muscle myosin, Effects of light chain phosphorylation and ionic strength*. J. Biol, Chem. 259, 8564-8571

22. Breckenridge, M. T., Dulyaninova, N. G., and Egelhoff, T. T. (2009) *Multiple regulation steps control mammalian non-muscle myosin II assembly in live cells.* Mol. Biol. Cell 20, 338-347

23. Somlyo, A. P., and Somlyo, A. V. (2003) *Ca2+ sensitivity of smooth muscle and non-muscle myosin II: modulated by G proteins, kinases, and myosin phosphatase.* Physiol. Rev. 83, 1325-1358

24. Lehrer S. S.; Morris E. P. (1984) *Comparison of the effects of smooth and skeletal tropomyosin on skeletal actomyosin subfragment 1 ATPase*. J. Biol. Chem. 259, 2070–2072.

25. Ostap E. M. (2008) *Tropomyosins as discriminators of myosin function*. Adv. Exp. Med. Biol. 644, 273–282.

26. Ma L, Bajic VB, Zhang Z. *On the classification of long non-coding RNAs*. RNA Biol. 2013;10(6):925-33.

27. Mercer TR, Dinger ME, Mattick JS. *Long non-coding RNAs: insights into functions.* Nat Rev Genet. 2009;10:155–9. doi: 10.1038/nrg2521

28. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. *Non-coding RNAs: regulators of disease*. J Pathol. 2010;220:126–39. doi: 10.1002/path.2638

29. Michael J. Milligan, Leonard Lipovich, *Pseudogene-derived lncRNAs: emerging regulators of gene expression*, Front Genet. 2014; 5: 476. Published online 2015 Feb 4. doi: 10.3389/fgene.2014.00476 PMCID: PMC4316772

30. Kim KK, Kim YC, Adelstein RS, Kawamoto S. *Fox-3 and PSF interact to activate neural cell-specific alternative splicing.* Nucleic Acids Res. 2010;39(8):3064-78.

31. Mironov, A. A., Fickett, J. W., and Gelfand, M. S. (1999). *Frequent alternative splicing of human genes*. Genome Res. 9. 1288-1293.

32. Modrek, B., Resch, A., Grasso, C., and Lee, C, (2001) *Genome wide detection of alternative splicing in expressed sequences of human genes.* Nucleic Acids Res, 29. 2850-2859.Myosins: a diverse superfamily.Sellers JR Biochim Biophys Acta. 2000 Mar 17; 1496(1):3-22
33. Luo T, Robinson DN. *The role of the actin cytoskeleton in mechanosensation. In: Kamkin A, Kiseleva I, editors. Mechanosensitivity in Cells and Tissues: Mechanosensitivity and Mechanotransduction.* Vol. 4. New York: Springer-Verlag; 2011. pp. 25–65

34. Bernstein E, Caudy A, Hammond S, Hannon G (2001). *"Role for a bidentate ribonuclease in the initiation step of RNA interference".* Nature. 409 (6818): 363–6. doi:10.1038/35053110. PMID 11201747

35. Saiki, R.; Gelfand, D.; Stoffel, S.; Scharf, S.; Higuchi, R.; Horn, G.; Mullis, K.; Erlich, H. (1988). *"Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase"*. Science. 239 (4839): 487–491. doi:10.1126/science.2448875. PMID 2448875

36. Salehi, Saeede & Naser Taheri, Mohammad & Azarpira, Negar & Zare, Abdolhossein & Behzad Behbahani, Abbas. (2017). State *of The Art Technologies Used to Explore Long Non-coding RNAs in Cancer Disease.* Cellular and Molecular of Medicine. 21. 10.1111/jcmm.13238.

37. Mohammadreza Hajjari, Abbas Salavaty(2015). *HOTAIR: an oncogenic long non-coding RNA in different cancers.Cancer Biology And Medicine.*12:1:15 doi: 10.7497/j.issn.2095-3941.2015.0006

38. Hua, Junjie & Ahmed, Musaddeque & Guo, Haiyang & Zhang, Yuzhe & Chen, Sujun & Soares, Fraser & Lu, Jennifer & Zhou, Stanley & Wang, Miranda & Li, Hui & Larson, Nicholas & K. McDonnell, Shannon & S. Patel, Parasvi & Yao, Cindy & Van der Kwast, Theodorus & Lupien, Mathieu & Y. Feng, Felix & Zoubeidi, Amina & Tsao, Ming-Sound & He, Housheng. (2018). *Risk SNP-Mediated Promoter-Enhancer Switching Drives Prostate Cancer through lncRNA PCAT19.* Cell. 174. 10.1016/j.cell.2018.06.014.

39. Hajjari M, Salavaty A. *HOTAIR: an oncogenic long non-coding RNA in different cancers*. Cancer Biol Med. 2015;12(1):1-9.

40. Petherick A (Aug 2008). "*Genetics: The production line*". Nature. 454 (7208): 1042–5. doi:10.1038/4541042a. PMID 18756228

41. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY (Jun 2007). *"Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs"*. Cell. 129 (7):1311–23. doi: 10.1016/j.cell.2007.05.022. PMC 2084369. PMID 17604720

42. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY (Aug 2010). "*Long noncoding RNA as modular scaffold of histone modification complexes"*. Science. 329 (5992): 689–93. doi:10.1126/science.1192002. PMC 2967777. PMID 20616235

43. Elsayed, E. T., Salem, P. E., Darwish, A. M., & Fayed, H. M. (2018). Plasma *long non-coding RNA HOTAIR as a potential biomarker for gastric cancer.* The International Journal of Biological Markers, 33(4), 528–533. https://doi.org/10.1177/1724600818760244
44. Lu, Xinyang; Liu, Zhiqiang; Ning, Xiaofei; Huang, Lunhua; Jiang, Biao, *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics*, Volume 26, Number 3, 2018, pp. 473-481(9). https://doi.org/10.3727/096504017X15105708598531

45. Loewen G, Jayawickramarajah J, Zhuo Y, Shan B. *Functions of lncRNA HOTAIR in lung cancer.* J Hematol Oncol. 2014;7:90. Published 2014 Dec 10. doi:10.1186/s13045-014-0090-4

46. Yi-Ting Chang,Tzu-Ping Lin,Jui-Ting Tang,Mel Campbell,Yun-Li Luo,Shih-Yen Lu,Chia-Pei Yang,Ting-Yu Cheng,Ching-Hsin Chang,Tze-Tze Liu,Chi-Hung Lin,Hsing-Jein Kung,Chin-Chen Pan,Pei-Ching Chang, *HOTAIR is a REST-regulated lncRNA that promotes neuroendocrine differentiation in castration resistant prostate cancer*, Cancer Lett. 2018 Oct 1;433:43-52. doi: 10.1016/j.canlet.2018.06.029

47. Shuai Xu, Jie Guo, Wei Zhang, *lncRNA PCAT19 promotes the proliferation of laryngocarcinoma cells via modulation of the miR-182/PDK4 axis.* Journal of Cellular Biochemistry. https://doi.org/10.1002/jcb.28552

48. Junjie Tony Hua, Musaddeque Ahmed, Haiyang Guo, Stephen N.Thibodeau, Paul C. Boutros, Housheng Hansen He, *Risk SNP-Mediated Promoter-Enhancer Switching Drives Prostrate Cancer through lncRNA PCAT19*, Cell. 2018 Jul 26;174(3):564-575.e18. doi: 10.1016/j.cell.2018.06.014

49. Zhang L, Geng Z, Meng X, Meng F, Wang L. *Screening for key lncRNAs in the progression of gallbladder cancer using bioinformatics analyses*. Mol Med Rep. 2018;17(5):6449-6455.

50. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. *The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression.* Genome Res. 2012;22:1775–89. doi: 10.1101/gr.132159.111.

51. Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, et al. *Transcriptome sequencing across a prostate cancer cohort identifies* PCAT-1*, an unannotated lincRNA implicated in disease progression.* Nat Biotechnol. 2011;29:742–9. doi: 10.1038/nbt.1914.

52. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. *Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript.* Nature. 2007;445:666–70. doi: 10.1038/nature05519

53. Blume SW, Meng Z, Shrestha K, Snyder RC, Emanuel PD. *The 5′-untranslated RNA of the human* dhfr*minor transcript alters transcription pre-initiation complex assembly at the major (core) promoter*. J Cell Biochem. 2003;88:165–80. doi: 10.1002/jcb.10326

54. Maenner S, Blaud M, Fouillen L, Savoye A, Marchand V, Dubois A, et al. *2-D structure of the A region of* Xist *RNA and its implication for PRC2 association.* PLoS Biol. 2010; 8:e1000276. doi: 10.1371/journal.pbio.1000276

55. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. *Long noncoding RNA as modular scaffold of histone modification complexes.* Science. 2010; 329:689–93. doi: 10.1126/science.1192002
56. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. *Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions.* Mol Cell. 2011;44:667–78. doi: 10.1016/j.molcel.2011.08.027

57. Nguyen VT, Kiss T, Michels AA, Bensaude O. 7SK *small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes.* Nature. 2001; 414:322–5. doi: 10.1038/35104581

58. Espinoza CA, Allen TA, Hieb AR, Kugel JF, Goodrich JA. B2 *RNA binds directly to RNA polymerase II to repress transcript synthesis.* Nat Struct Mol Biol. 2004; 11:822–9. doi: 10.1038/nsmb812

59. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. *The nuclear-retained noncoding RNA* MALAT1 *regulates alternative splicing by modulating SR splicing factor phosphorylation.* Mol Cell. 2010; 39:925–38. doi: 10.1016/j.molcel.2010.08.011

60. Tsuiji H, Yoshimoto R, Hasegawa Y, Furuno M, Yoshida M, Nakagawa S. *Competition between a noncoding exon and introns:* Gomafu *contains tandem UACUAAC repeats and associates with splicing factor-1*. Genes Cells. 2011; 16:479–90. doi: 10.1111/j.1365-2443.2011.01502.

61. Muddashetty R, Khanam T, Kondrashov A, Bundman M, Iacoangeli A, Kremerskothen J, et al. *Poly(A)-binding protein is associated with neuronal BC1 and BC200 ribonucleoprotein particles*. J Mol Biol. 2002; 321:433–45. doi: 10.1016/S0022-2836(02)00655-1

62. Lin D, Pestova TV, Hellen CU, Tiedge H. *Translational control by a small RNA: dendritic* BC1 *RNA targets the eukaryotic initiation factor 4A helicase mechanism*. Mol Cell Biol. 2008;28:3008–19. doi: 10.1128/MCB.01800-07

63. Watanabe Y, Yamamoto MS. *S. pombe mei2+ encodes an RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species meiRNA.* Cell. 1994; 78:487–98. doi: 10.1016/0092-8674(94)90426-X

64. Campalans A, Kondorosi A, Crespi M. *Enod40*, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*. Plant Cell. 2004; 16:1047–59. doi: 10.1105/tpc.019406

65. Feng J, Funk WD, Wang SS, Weinrich SL, Avilion AA, Chiu CP, et al. *The RNA component of human telomerase.* Science. 1995;269:1236–41. doi: 10.1126/science.7544491.

66. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. *Functional demarcation of active and silent chromatin domains in human* HOX *loci by noncoding RNAs.* Cell. 2007;129:1311–23. doi: 10.1016/j.cell.2007.05.022

67. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. *Long noncoding RNA as modular scaffold of histone modification complexes.* Science. 2010;329:689–93. doi: 10.1126/science.1192002

68. Ying Qu, Bingchen Han, Yi Yu, Weiwu Yao, Shikha Bose, Beth Y. Karlan, Armando E. Giuliano, Xiaojiang Cui, *Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells*, https://doi.org/10.1371/journal.pone.0131285

69. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, et al. *ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome*. Cell Stem Cell. 2007;1(5):555–67. pmid:18371393; PubMed Central PMCID: PMC2423808

70. Liu S, Cong Y, Wang D, Sun Y, Deng L, Liu Y, et al. *Breast Cancer Stem Cells Transition between Epithelial and Mesenchymal States Reflective of their Normal Counterparts*. Stem Cell Reports. 2014;2(1):78–91. Epub 2014/02/11. S2213-6711(13)00149-5 [pii]. pmid: 24511467; PubMed Central PMCID: PMC3916760.

71. Gest C, et al: *Rac3 induces a molecular pathway triggering breast cancer cell aggressiveness: differences in MDA-MB-231 and MCF-7 breast cancer cell lines.* BMC Cancer 13: 63, 2013

72. D'Anselmi F, Masiello MG, Cucina A, Proietti S, Dinicola S, Pasqualato A, Ricci G,Dobrowolny G, Catizo ne A, Palombo A, Bizzarri M: *Microenvironment promotes tumor cell reprogramming in human breast cancer cell lines.* PLoS One 8:e83770, 2013

73. Pérez-Yépez EA, Ayala-Sumuano J-T, Reveles-Espinoza AM, Meza I: *Selection of a MCF-7 Breast Cancer Cell Subpopulation with High Sensitivity to IL-1β: Characterization of and Correlation between Morphological and Molecular Changes Leading to Increased Invasiveness*. Int J Breast Cancer 2012:609148, 2012

74. Zhang, H.et al. *Inhibited Indirectly by LincRNA HOTAIR, Directly Inhibits SETDB1 and Reverses the EMT of Breast Cancer Stem Cells by Downregulating the STAT3 Pathway*. Stem Cells, 32: 2858-2868. doi:10.1002/stem.1795

75. Sørensen, K.P., Thomassen, M., Tan, Q. et al. *Breast Cancer Res Treat* (2013) 142: 529. https://doi.org/10.1007/s10549-013-2776-7

76. Tano K, Akimitsu N. *Long non-coding RNAs in cancer progression*. Front Genet. 2012; 3:219. Published 2012 Oct 24. doi:10.3389/fgene.2012.00219.

77. Rui-Xue Tang, Wen-Jie Chen, Rong-Quan He, Jiang-Hui Zeng, Liang Liang, Shi-Kang Li, Jie Ma, Dian-Zhong Luo and Gang Chen, China, *Identification of a RNA-Seq based prognostic signature with five*

lncRNAs for lung squamous cell carcinoma, Oncotarget. 2017; 8:50761 50773. https://doi.org/10.18632/oncotarget.17098

78. Litwa, Karen & Horwitz, Rick & Lamers, Marcelo. (2015). Non-Muscle myosin II in disease: *Mechanisms and therapeutic opportunities.* Disease models & mechanisms. 8. 10.1242/dmm.022103.

79. Tano, K., & Akimitsu, N. (2012). *Long non-coding RNAs in cancer progression*. Frontiers in genetics, *3*, 219. doi:10.3389/fgene.2012.00219