Changes in microRNA target sites attributed to single nucleotide polymorphisms may influence breast cancer susceptibility

A THESIS SUBMITTED

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE
IN
LIFE SCIENCE

SUBMITTED TO

NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA

BY

Anshu Kumari

Roll no: 412LS2040

UNDER THE SUPERVISION OF

DR. BIBEKANAND MALLICK
ASSISTANT PROFESSOR

DEPARTMENT OF LIFE SCIENCE
NATIONAL INSTITUTE OF TECHNOLOGY
ROURKELA-769 008, ODISHA, INDIA

राष्ट्रीय प्रौद्योगिकी संस्थान, राउरकेला
National Institute of Technology, Rourkela

Dr. Bibekanand Mallick, M.Tech., Ph.D.
Assistant Professor

RNAi & Functional Genomics Lab.
Department of Life Science
National Institute of Technology
Rourkela - 769 008, Odisha, India
Telephone: +91-661-246 2685 (O)
E-mails: vivek.iitian@gmail.com,
mallickb@nitrkl.ac.in
Website: http://vvekslab.in

CERTIFICATE

*This is to certify that the thesis entitled "**Changes in microRNA target sites attributed to single nucleotide polymorphisms may influence breast cancer susceptibility**" submitted by **Ms. Anshu Kumari** (Roll No: 412LS2040) in partial fulfilment of the requirements for the award of **Master of Science in Life Science** to the National Institute of Technology, Rourkela, is an authentic and original record of research work carried out by her under my supervision and guidance.*

To the best of my knowledge, the work incorporated in this thesis has not been submitted elsewhere for the award of any degree.

10/05/2014

Place: Rourkela

Date: 10.05.2014

(Dr. Bibekanand Mallick)

Dr. Bibekanand Mallick
Assistant Professor
Department of Life Science
NATIONAL INSTITUTE OF TECHNOLOGY
Rourkela-769008, Odisha

ACKNOWLEDGEMENTS

As the knowledge is infinite so is the greatness to teacher, who passes the knowledge and experience to the best of his capabilities to his student. I find myself fortunate enough to express my sincere and profound sense of gratitude to my learned and respected guide Dr. Bibekanand Mallick, NIT Rourkela, Orissa for his expert guidance and constant inspiration throughout the work that paved the way for successful completion of this endeavor. I don't have appropriate words in vocabulary to express my deep sense of ever indebtedness to him.

I am very grateful to Dr. Samir Kumar Patra, H.O.D, Department of Life Science, NIT Rourkela, who encouraged me throughout my work. I also sincerely thank my teachers Dr. Rasu Jaybalan, Dr. Sujit Kumar Bhutia, Dr. Surajit Das and Dr. Suman Jha for their untiring help, expert guidance and critical comments to make difficult tasks simple.

My deepest gratitude goes to Mr. Bedanta B. Mohanty, Debashree Das, Devyani Samantarrai, Mousumi Sahu, Garima Singh, for their kindly and timely cooperation and encouragement during the entire tenure of my work. They stood beside me through the thick and thins of present work. Without their cooperation, it would have been impossible for me to complete the work.

Finally, I deeply express my gratitude to my parents, my family and Ashwani Kumar for providing their unconditional emotional support and love.

Date: 11 May, 2014 Anshu Kumari

Place: Rourkela

List of tables:

List of figures:

Contents:

Abstract

MicroRNAs (miRNAs) are a class of small, ~22-nt-long endogenous, single-stranded non-coding RNAs (ncRNAs) that takes part in post-transcriptional gene silencing through deadenylation, translational repression, and destroying their target messenger RNAs (mRNAs) They exert their regulatory role on the corresponding mRNAs by binding to their 3'UTR, CDS and 5'UTR. When SNP occurs in 3' UTRs, it changes the mRNA stability and translation by altering the miRNA::mRNA interactions. SNPs present in the 3'UTR of genes associated with breast cancer might contribute to the initiation and susceptibility of the malignancy. In this work, we aim to find the significant pairs of target SNP:miRNA by using the different *in silico* methods. From our analysis, we screened and got six potential target SNP pairs which have been proposed to be involved in breast cancer risk in females.

Key word: miRNA, SNP, UTR, CDS, mRNA, non-coding RNA.

Introduction

Introduction:

Cancer is a class of deadly disease characterized by uncontrolled cell growth. More than 200 different types of cancer have been identified till now. Each cancer is classified on the basis of type of cell that is initially affected. When growth of the damaged cell is uncontrolled, it forms lump or masses of tissue, known as tumors. Tumor cells release certain chemicals (cytokines, hormones) that alter normal function of the body. A benign tumor like skin wart is always present in its original location and do not invade or spread to the surrounding tissues. A malignant tumor, however invades and spread to the surrounding tissues through the circulatory system and lymphatic system. These malignant tumors are known as cancer.

Cancer possesses several biological capabilities that are acquired during the development of human tumors. These hallmarks basically constitute an organizing principle for rationalizing the complexities of the cancer. These are tissue invasion, genetic instability, tumor promoting inflammation, sustained growth signaling, evading anti-growth signaling, resisting programmed cell death, enables repilcative immortality, deregulated cellular metabolism, avoiding immune destruction, inducing new blood flow (Hanahan and Weinberg, 2011).

Figure 1: Hallmarks of cancer.

Major causes of cancer are different lifestyle, genetic disorder, carcinogens, virus, and bacterial infection. Carcinogens are cancer causing agent or substance which damaged the DNA and promoting the cancer. For example tobacco, gamma ray. Different life style like Smoking, handling of toxic chemical, high fat diet is also factor for generating cancer. Some genetic disorder like beckwith-wiedemann syndrome and wiskott-Aldrich alter the normal immune system of the body and cause cancer. Cancer may also be caused by the infection of certain virus for example HIV and Epstein-Barr virus. HIV that causes AIDS also developing some type of childhood cancer.

There are many different types of cancer. One type of cancer is breast cancer. Breast cancer is the most common type of cancer in females. Breast cancer generally starts from the inner lining of milk duct. It can also start from the lobules that supply milk. A breast cancer that began from lobules is known lobular carcinoma, other one which is developed from the lactiferous ducts is known as ductal carcinoma. Lobular carcinoma is very less common type of cancer but ductal carcinoma is most common type of breast cancer.

Breast cancer (BC) is the most common type of cancer in female, every year 1 million or more then1 million new cases diagnosed. Current data of National cancer Institute in USA, 232,340 female breast cancers reported per year and 2,240 male breast cancers reported per year. 39,630 deaths reported per year. It includes both males and females. Ratio of breast cancer patients are very high in developed countries in comparison to small developing ones. This cancer is generally found in elderly aged women. Different life styles and eating habit of females are some of the important factors contributing to breast cancer. Though the exact causes of breast cancer are not known but various genetic and epigenetic changes are the ultimate focal point of oncogenesis in breast cancer.

Epigenetic regulations, which regulate the expression of genes without changing its sequence, forms major part of etiology in breast cancer. One of the epigenetic changes is RNA interference. The RNA interference (RNAi) heralded a revolution in RNA biology which contributes significantly in regulating the expression of gene epigenetically. Small non-coding RNAs, eg. miRNA, siRNA, piRNA etc, comprises a greater part in RNA interference mechanism (Siomi and Siomi, 2009). Some of the miRNA involve in the breast cancer development. Micro

RNAs (miRNAs) are a class of small, ~22-nt-long endogenous, single-stranded non-coding RNAs (ncRNAs) that takes part in post-transcriptional gene silencing through deadenylation, translational repression, and destroying of their target messenger RNAs (mRNAs) (Mallick and Ghosh, 2012).

The oncogenesis process is also affected by polymorphism. A great example of polymorphism is human ABO blood group and Rh factor. Single nucleotide polymorphisms (SNPs) are variation occurring in single nucleotide base on a DNA sequence. Generally 2 types of SNPs are present, synonymous and non-synonymous. SNP that change the function of protein by changing the amino acid sequence is called as non-synonymous SNP. Synonymous SNP do not change the protein sequence. It is also called as silent mutation.SNP may occur in the coding region as well as non-coding region of a gene. These SNPs may affect whether and how organism develops certain disease and responds to different chemicals and drugs. SNP have great value in biological research and drug development and provide details of person's susceptibility. (Mahdi et al., 2013)

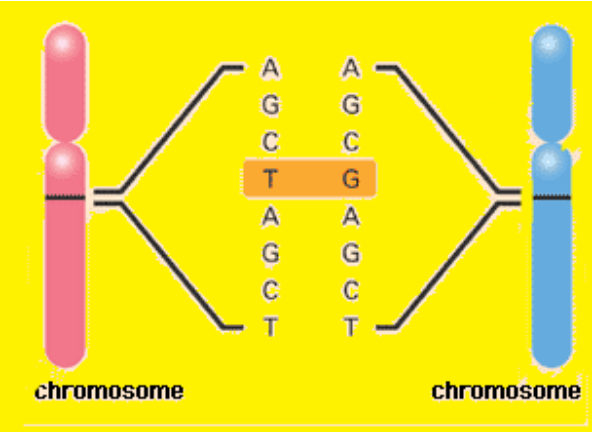

Figure 2: Sequence showing single nucleotide polymorphism (SNP)

Some of the SNPs are found in miRNA or miRNA regulated gene. These can alter or change the miRNA binding sites. We know that miRNA modulate post-transcriptional gene regulation. miRNAs bind with protein coding gene (PCG) fully or partially with its

complementary sequence in 3’ UTR region. Sometime miRNA also bind with the coding sequence and 5’UTR of the mRNA. When miRNA bind with mRNA, it changes the stability and translation mechanism of mRNA. When SNP occurs in 3’ UTRs, it changes the mRNA stability and translation. It alters protein::mRNA, miRNA::mRNA interaction. SNPs present in the 3’UTR disrupt the miRNA targeting which might contribute to cancer initiation and susceptibility. Based on this immense source of information, it can now be possible to uncover the etiological heterogeneity, clinical course, and response to treatment of different human cancers.

miRNAs and single nucleotide polymorphisms associated with breast cancer have been collected from various literatures and online databases (dbSNP, miRBase, PhenomiR etc). Structural features of the 3’ UTR region of the mRNA have been studied by using different online tools (RNAfold, mfold etc). miRNA:mRNA interaction study has been carried out using customized perl scripts to find the interaction characteristics of miRNA to target 3′UTR with and without SNPs (nature of binding, binding energy, binding potential etc). Resulted pairs have been filtered and most significant pair has been selected for further validation.

Review of literature

Review of literature:

RNA has been discovered in the late 1800s, but importance of RNA in the cell is highlighted after DNA and protein. DNA is copied into the mRNA, then it is translated into the protein by the help of other two RNAs- tRNA and rRNA. In the recent years, the study of genome-wide approach and genome sequencing provided the details of mammalian transcriptome. It was revealed that mammalian transcriptome includes large number of coding and non-coding sequences. Coding sequence or Exon region code for protein. Non-coding RNA are RNA that do not code for any protein. Two types of non-coding RNAs are present, i.e. small non-coding RNA and long non-coding RNA. These small non-coding RNA include functionally important and highly abundant RNAs such as snoRNA, miRNA, piRNA, siRNA, snRNA, tiRNA. These RNAs are involved in many cellular processes, such as epigenetic regulation, transcriptional regulation, post-transcriptional regulation etc.

Victor Ambros and colleagues discovered a gene lin-4, this lin-4 control the larval development in *C. elegans* and does not code for any protein. These lin-4 RNAs are double stranded and have sense and antisense strand. The antisense strand of lin-4 bind with multiple sites on 3'UTR of lin-14 gene. This complementary binding of lin-4 to the lin-4 gene reduces the lin-14 protein. This lin-4 recognized as the regulatory RNAs called as microRNA (miRNA) (Bartel, 2004).

miRNA is a small, 22 nucleotide long endogenous, single stranded non-coding RNA. It takes part in posttranscriptional gene silencing through translational repression and deadenylation of their target site. miRNAs play major role in different biological processes in plant and animal such as differentiation, organogenesis. tumorigenesis, cell proliferation, apoptosis, and embryogenesis. The aberrant expression and deregulation of miRNA are implicated in risk of several diseases such as cancer. miRNA bind with 3'UTR of the mRNAs. Some time, it also binds with the CDS and 5'UTR. More than 30% human genes are targeted by the miRNAs.

Biogenesis of miRNA:

miRNA are basically encoded in protein coding genes or intergenic regions. The transcription of intergenic miRNA from primary miRNA (pri-miRNA) takes place by the enzymatic activity of RNA polymerase III and RNA polymerase II. Intronic miRNA are co-transcribed with their host gene. Several enzymes and proteins are involved in formation of miRNA. Two RNase III enzymes called Drosha and its regulatory subunit DGCR8 take part in cleavage of pri-miRNA transcript and generate a hairpin precursor called pre-miRNA in the nucleus. Pri-miRNA hairpin stem contain 33 base pairs with terminal loop and upstream and downstream of the hairpin. DROSA cleave 11 base pair away from base of hairpin. This hairpin may be single stranded or double stranded RNA junction. Hairpin stem act as a recognition element for binding of DGCR8. It determines the cleavage site for Drosha processing. Sometime Pri-miRNA loops also act as a binding site for nuclear ribonucleoprotein A1 (hnRNP A1). It change in conformation of hairpin create more favorable cleavage site for Drosha processing. (Mallick and Ghosh, 2012) After act of Drosha, pre-miRNA is then exported to the cytoplasm from nucleus by expotin-5-RanGTP. It is a nuclear transport receptor complex. Dicer acts as catalysis for formation multiple class of small non-coding RNA. In the mammals system, Dicer is aided by trans- activation responsive RNA-binding protein also called as TRBP for special function. In the cytoplasm dicer processed this pre-miRNA into mature 22 nt miRNA. Dicer and some accessory protein act on pre-miRNA and complete the process of miRNA biogenesis.

In the next step of miRNA production, pre-miRNA is associated with Argonaute protein family. During the loading process, non-guided strand is cleaved by the Argonaute protein. Argonaute protein family play major role in gene silencing pathway. Ago protein is highly conserved in between species.

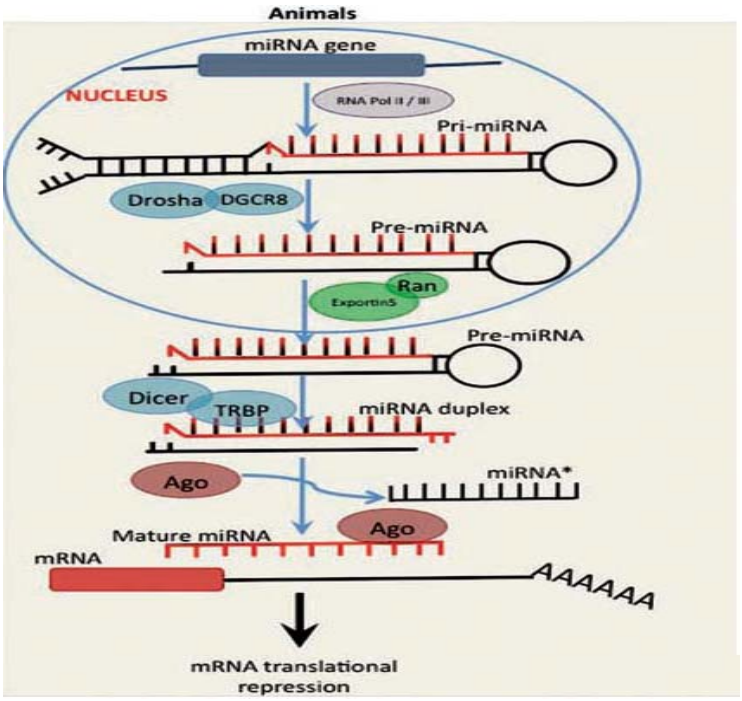

Figure 3: Step of biogenesis of miRNA

miRNA targeting:

miRNA target the mRNA and direct mRNA destabilization, and translational repression. So the question arise how this tinny miRNA takes part in regulation and find their target site. miRNA target recognition is observed from the finding on lin-4. Lin-4 had several sequences which is complimentary to the 3' UTR of lin-14 mRNA (Bartel, 2009). 2-7 nucleotides of miRNA that bind with mRNA is called the Seed region. This binding is Watson and crick base pairing. Seed region is important for miRNA target recognition.

Facts/Factors of Target binding:

There are several factors that determine efficient miRNA binding.

1. miRNA bind with the 3'UTRs of the mRNA, but not bind to the two end of 3'UTR.
2. Position of miRNA binding in 3'UTR should at least 15 nt away from the stop codon.

3. Position should away from the center of long UTR.
4. If one miRNA bind multiple site so this miRNA will have great effect in regulation.
5. A-U rich nt composition near the target site.
6. Structure of mRNA should be feasible their for binding of the miRNA.

Types of target site:

1. Canonical site.
2. Marginal site.
3. Atypical site.

Canonical site- It is most traditional and well know site for miRNA binding. Three type of canonical site is present,

- 7 mer-A1 site
- 7 mer-M8
- 8 mer

7mer-A1 site binding will be from 2-7 nucleotides from the $5^{/}$ end of miRNA and "A" will be in target mRNA corresponding to first position of miRNA. In 7mer-m8, binding from 2 to 8^{th} position of miRNA occurs. In 8mer, 2-8 base pairing is present plus "A" on the target mRNA corresponding to first position of miRNA. Second type is marginal site consisting of two type of sites- one is 6 mer (2-7 base pairing) and another is offset 6 mer (3-8 base pairing).

Third is Atypical site. It also contains two types:

1. 3'Supplementary site
2. 3'compensatory pairing.

By certain experimental evidence proved that 3'supplementary pairing increase the binding specificity and affinity of seed pairing. 3' supplementary pairing is productive and associated with different sufficient number of sites. Beside the conventional seed region complementarily

additional base pairing has been found at 12-17 nt position of mRNA know as 3' supplementary site. Compensatory site, Watson-crick pairing usually centering on miRNA nucleotide 13-16 can compensate for seed mismatch and thereby create a functional site. If a seed region has some of the wobbles pairing the pairing compensate by the compensatory pairing. Although miRNA has diverse binding sites but 8mer binding site has been found to be most efficient one for successful target degradation.

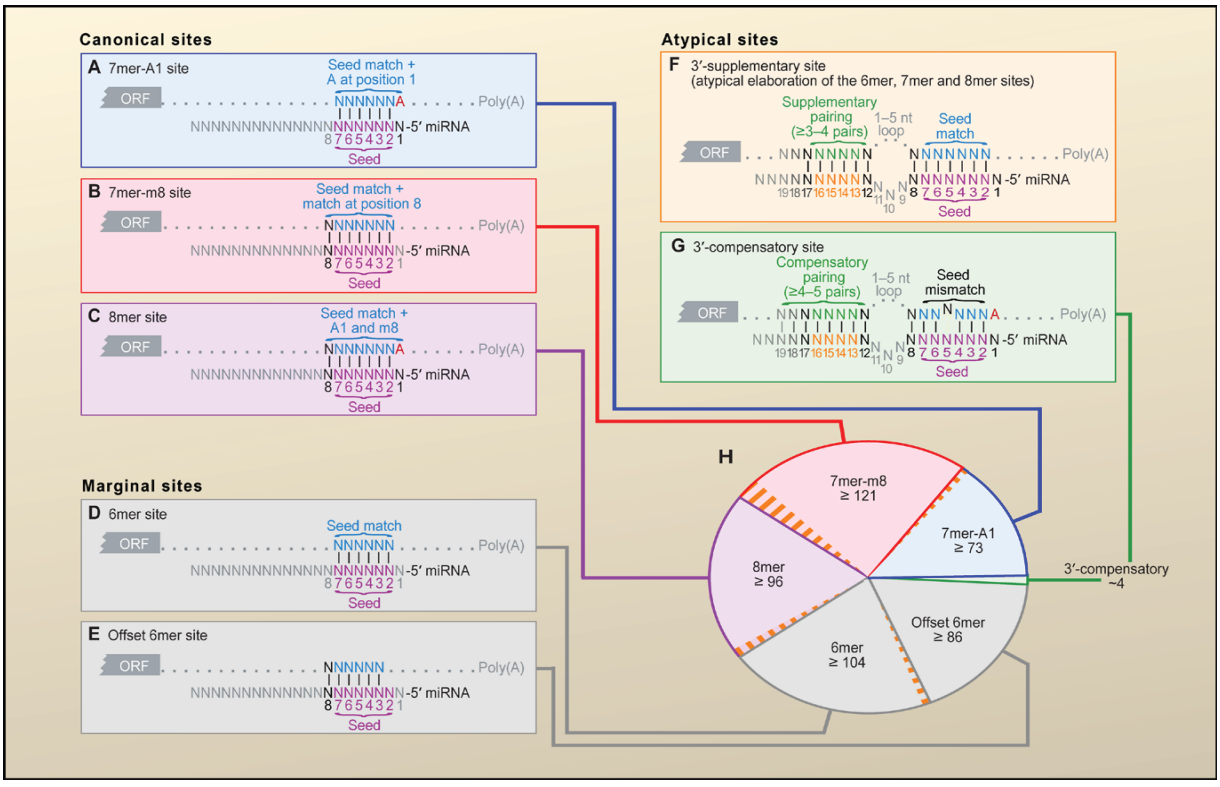

Figure 4: Different type of target site (Bartel, 2009)

Polymorphism defined as a different form of an allele in a gene. It defines as morphs of the phenotype. More than 1% genetic variation in a population called as Single nucleotide polymorphism . The SNP is common single nucleotide variation in the human genome. It is a type of mutation in an individual, but the percentage of the mutation rate is high. SNP arises because of point mutation. The SNP may present in the coding region, non-coding region or intergenic regions. Most of SNP occur in non-coding region.(Gray et al., 2000)

There are two types of base substitution resulting in SNPs

- **Transition substitution**- In this substitution change in nucleotide happen in between purines to purines and pyrimidines to pyrimidines.
- **Transversion substitution**- change of nucleotide occurs in between purine and pyrimidines.

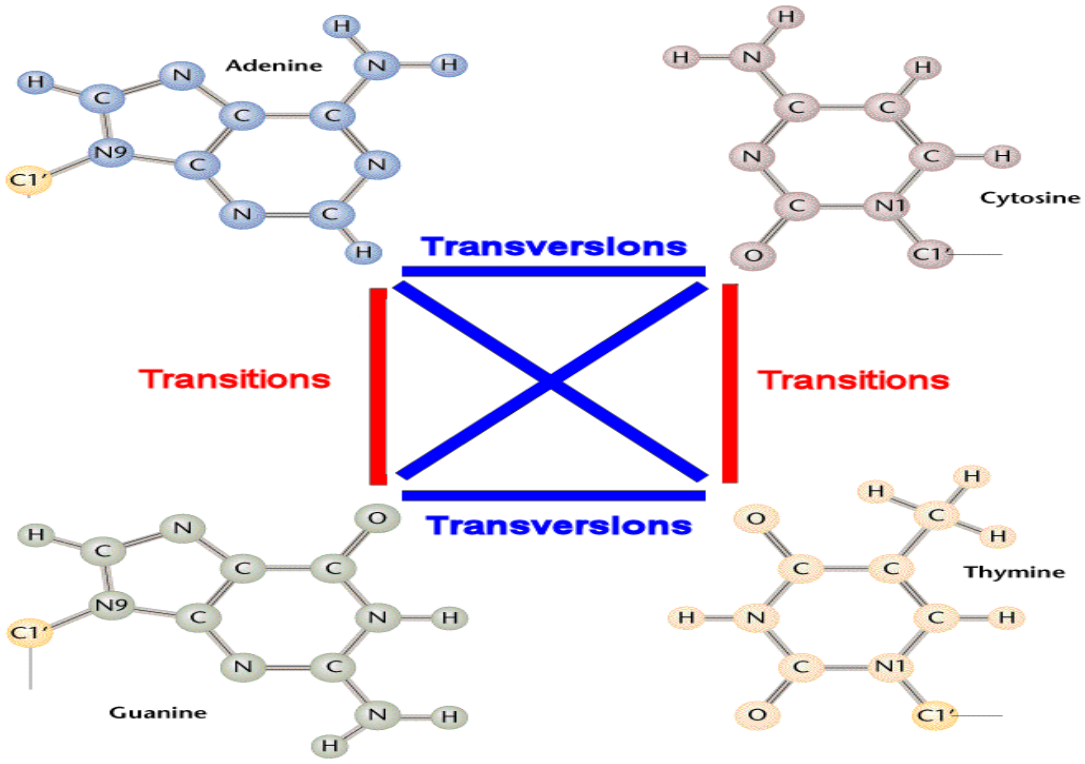

Figure 5: Two types of base substitution

Distribution of SNPs

SNP presents in entire human genome. SNP rate is very lower in sex chromosome. Most of the SNP is present in the coding region.

Synonymous: Addition on deletion of nucleotide not change the amino acid, so no change in the protein. It is also called as silent mutation.

Non- Synonymous: Addition on deletion of nucleotide changes the amino acid. It is missense mutation. It changes the protein. Most of the SNP result in non-synonymous codon changes.

Importance of SNPs:

- It is used to find the genetic similarity and dissimilarity among the population.
- Finding the disease gene.
- SNP detection explains and diagnoses many diseases.
- Study variation in drug responses.

Role of SNPs in cancer:

Genetic variation has silent, harmful & harmless effects. By the help of "SNP" now people can observe different level of individual cancer risk. This type of variation found in the coding region. SNP changes the amino acid and because of changes in one amino acid 3D structure of protein is changed. This protein changes the metabolism of the body and makes the person susceptible to cancer. This protein converts the procarcinogen into the carcinogen. In a normal person this protein is absent, but the person has SNP in their gene produced protein, which convert the procarcinogen substance to carcinogen substance very actively.

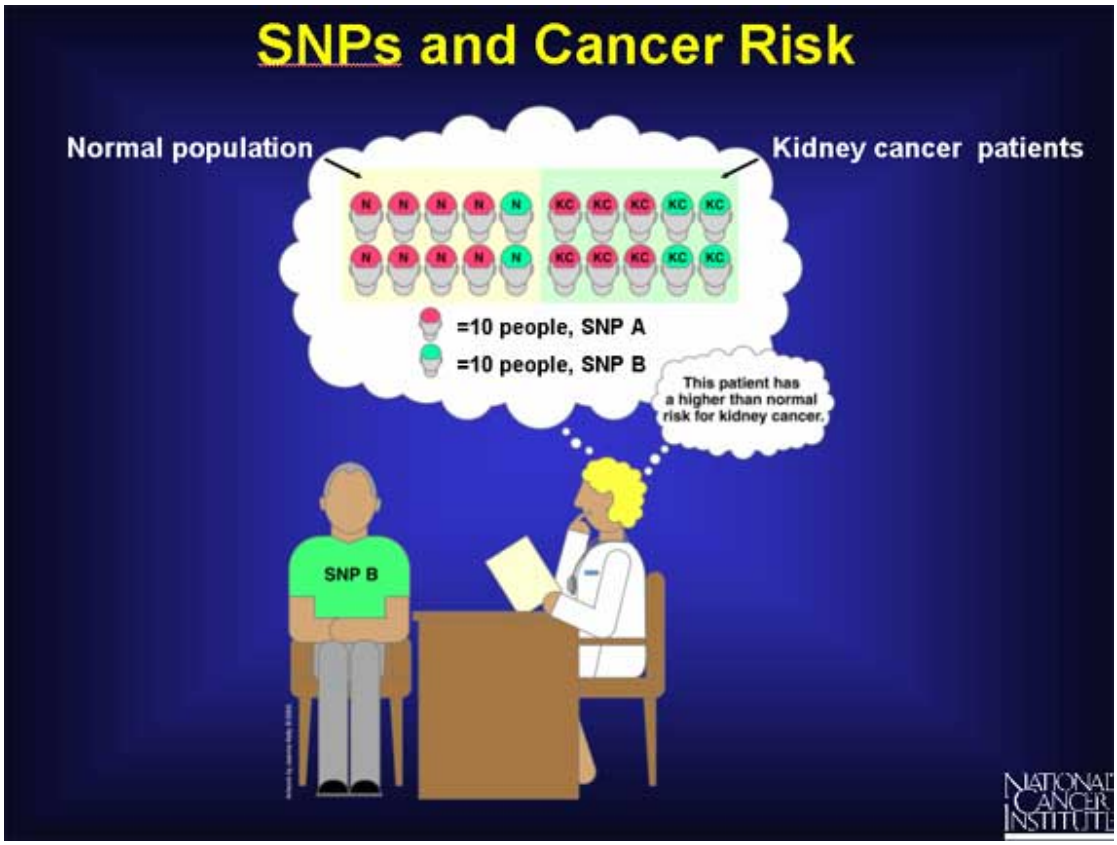

Figure 6: SNP and cancer risk.

SNP regulates the miRNA targeting is also one of the causes of different cancers. miRNAs have fundamental importance in growth control, differentiation and human disease such as cancer. miRNAs bind with the 3' UTR of mRNA. SNP affects the miRNAs binding. If the SNP present in a miRNA target site, then it implicated in several diseases. By the recent annotation study of SNP and miRNA gene give the detail of change of miRNA targeting because of SNP. (Stahlhut Espinosa and Slack, 2006). SNP present in seed region it alter the miRNA-mRNA targeting. A SNP can change, create and modified the miRNAs targeting site. This change decrease or increase the protein translation mechanism. When SNP present in 3' UTR gene it implicated in many cancer and enhanced the tumorogenisity. SNP effects on miRNAs targeting according to the difference in binding free energy and alignment scores. 3'UTR are likely to be downregulated by particular miRNA. (Xu et al., 2013).

SNP increases the breast cancer risk in female. Mutation in these genes such as ATM, BARD 1, PMS1, BARCA1, BARCA2, BRIP1, CDH1, APE1 increase the breast cancer. For example, the human DNA repair system prevents the DNA damage caused by certain environmental agent. AP endonuclease 1 (APE1) is one of the protein, which take part in the DNA repair mechanism. It acts as 3'-phosphodiesterase and repair DNA single strand breaks. It directly removes the damaged bases by enzyme action. It is a promising tool used for anticancer therapy. Polymorphism in DNA repair gene influence the susceptibility to carcinomas. Polymorphism in APE1 was associated with the several cancer risks. In human genome is APE1 located on chromosome 14q11.2-q12 and encodes a 317 amino acid protein. Several sequence variation present in this gene changes the amino acid from aspartic acid to glutamic acid(SNP id- rs3136820). Meta-analysis suggested that the APE 1 rs3136820 polymorphism was a major risk factor for breast cancer (Zhao et al., 2014). PPMID gene also involve in increase risk of breast cancer. If the women with PPMID gene mutation then she have 20% chance to developing breast cancer.(Mahdi et al., 2013). 5-10% of breast cancer are thought to be hereditary. That is caused by the abnormal gene. Some of the gene such as breast cancer susceptibility gene 1(BRCA1) and breast cancer susceptibility gene 2 (BRCA 2) are hereditary gene associated with breast cancer. Mutation in BRCA1 and BRCA2 gene in germline cell cause susceptibility to breast cancer. If the high frequency of mutation showed in gene so the gene product is changed. If a female carries mutated BRCA1 and BRCA2 gene, then she has high risk to develop breast cancer during lifetime. Checkpoint kinase2 (CHEK2) encode a G2 checkpoint kinase. It is involved in DNA damage repair. CHEK2 1100delC gene mutation increases the breast cancer.

SNP rs-13281615 located in the 8q24 chromosomal region involved in the development of many types of cancer including breast cancer. This SNP location is nearby myelocytomatosis oncogene (MYC) which promotes diffensiation , cell proliferation and transformation. Variation in gene alter the c-MYC expression and increased breast cancer susceptibility.(Gong et al., 2013) SNP in the miR367 binding site at the 3'UTR of the RYR3 gene affects the breast cancer risk. RYR3 is third isoform of the RYR family. It is calcium induced calcium release channel protein and controls the cytoplasm calcium level. RYR is commonly expressed in breast cancer. It is a

biomarker for identify breast cancer prognosis and oncogenic mutation. RYR3 plays a role in breast calcification. By Bioinformatics prediction provides the detail of SNP present in 3'UTR of RYR3 gene and associated with cancer risk. RYR3 gene is the principal regulator of breast cancer cell growth. Genetic variation in the RYR3 gene affects the intracellular calcium influx and cell proliferation as well. (Zhang et al., 2011).

Epistasis or gene-gene interaction has also contributed to increased breast cancer risk. Sometime singe SNP not able to Cause Breast cancer. SNP-SNP interaction is contributed to influencing breast cancer risk. (Onay et al., 2006).

For study of miRNA –SNP interaction various data is collected from different database. (Nicoloso et al., 2010). The human miRNA sequence is collected from miRbase and SNP collected from the NCBI site. Interact the collected sequence by use of Bioinformatics tool. Find those pairs which are energetically efficient and calculate the minimum free energy of this pairing. Find the energetically relevant pair. By several interactions studies, we conclude that the BC susceptibility increased because of the SNP mediated change, create and modified of miRNA targeting.

OBJECTIVES:

OBJECTIVE 1

Data retrieval: Collection of SNPs and miRNAs involved in breast cancer.

OBJECTIVE 2

Target prediction: Study of interaction between SNPs and miRNAs

OBJECTIVE 3

Identification of significant pairs of target SNPs-miRNAs

Materials and Methods

Materials and methods:

DATA retrieval:

Collection of SNP: Extensive literature studies have been carried out to find the single nucleotide polymorphisms reported to be associated with breast cancer from. SNPs present in different genic regions, ex. 3'UTR, 5'UTR, intron, CDS, others, have been retrieved. The sequence informations of the SNPs have been retrieved from dbSNP database.

miRNA data mining: Differentially expressed microRNAs in breast cancer system have been retrieved from various published literatures and online databases (ex. PhenomiR, OncomiR). Up regulated microRNas were taken for the interaction study. The sequence informations of the miRNAs have been retrieved from miRBase database.

Target prediction:

1. We have collected the ancestral and mutated SNP sequences in breast cancer system from dbSNP. The text file was splitted into three files separating ancestral sequence, mutated sequence and variation using customized pearl script.
2. The upregulated miRNAs in breast cancer obtained from various databases were considered for the miRNA:SNP interaction study for both the cases i.e., ancestral and mutated SNP dataset using RNAhybrid. The criteria set for the above interaction study are as follow.
    a) 8mer binding sites with binding energy < -1 kcal/mol
    b) 7mer binding sites with binding energy < -1 kcal/mol
    c) Off-set 6mer binding sites with binding energy < -1 kcal/mol
    d) Interactions with binding energy < -10 kcal/mol
3. For both the cases the above interaction files were merged into a single file and further filtration to find best interaction was carried out using customized pearl scripts basing on the binding pattern and binding energy.

4. R statistical package was used to sort the miRNA:SNP pair on the basis of percentage minimum free energy change between miRNA interaction with the ancestral and mutated sequences.
5. Finally, five interactions, which were significantly affecting breast cancer etiology have been filtered and the pairs were taken for further validation studies showing significant differences in binding pattern and binding energy.

Structural Analysis:

Structural analysis of the folded mRNA has been carried out to check the site availability for binding of miRNA and to compare the site morphology due to presence and absence of SNP.

- The sequence of the mRNA was retieved and taken for folding study.
- The sequence was folded using different online tools (ex, RNAfold, mfold).

Cell culture:

Cell culture is a process by which a cell grows under the control condition outside of the natural environment by proving media and certain growth regulator. MDA-MB-231 is human breast carcinoma cell line which was obtained from national center for cell science (NCCS), Pune. DMEM media with 10% FBS (Fetal bovine serum) was used for culturing the cell. Nunc^{TM} cell culture flask is used. The cell culture flask contain breast cancer cell line kept in CO_2 incubator. CO_2 level is maintained at 5%. Medium contained phenol red it defines the pH level in the medium. The incubator temperature set at 37^0C for maintaining the physiological temperature in the cell culture environment.

Steps of cell culture:

1. Cell Harvesting: 1x 10^7 cells were grown in suspension. Cells were pelleted down by centrifugation for 5 min at 300 x g. Suparnateant was removed carefully. Direct lysis in the vessel at the appropriate volume of buffer RLT.

2. Tryprinization of cell: Medium is removed and cell was washed with phosphate buffer saline (PBS). Then PBS is removed and 0.1-0.25% trypsin was added. Kept for 2-3 min. The cell was detached from the flask surface. After this cell culture medium was added. No need to remove trypsin from the cell. A medium containing the serum to inactive the trypisin. In next step cell was transferred to an RNase-free glass or polypropylene centrifuge tube and centrifuged at 300x g for 5 min. Supernantent was removed carefully. Cells are resuspend in fresh media and put in incubator.

Results and Discussions

Results and discussions:

Extensie literature studies have been carried out to single nucleotide polymorphisms reported to be present in different genic regions, ex. 3'UTR, 5'UTR, Exon, Intron, CDS etc. In breast cancer, 59 SNPs were reported in 3' UTR, 9 SNPs were reported in 5' UTR, 119 SNPs were reported in CDS, 242 SNPs were reported in Introns,175 SNPs were reported in other regions.

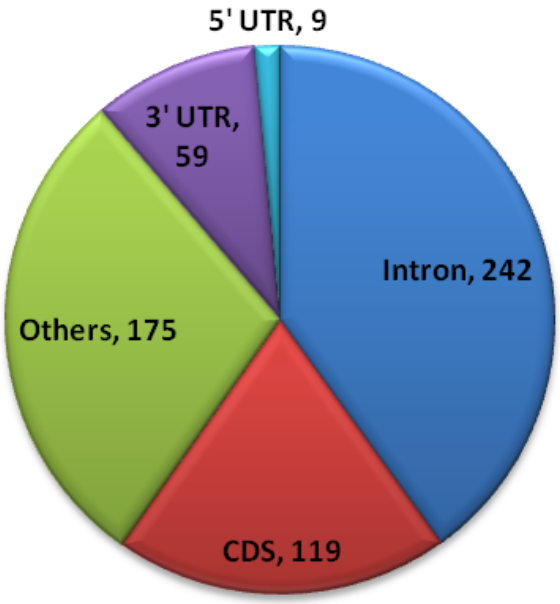

Figure 7: Distribution of SNP in different region of breast cancer associated gene

MicroRNAs that are differentially expressed in breast cancer have been retrieved from different published literatures and online databases(Ex miRBase, PhenomiR, miRCancer etc). 75 miRNAs were found to be upregulated in breast cancer system.

Table 1. Name of 76 upregulated miRNAs reported in breast cancer.

hsa-let-7d*	hsa-miR-30c-1*	hsa-miR-3197	hsa-miR-601
hsa-miR-1180	hsa-miR-30c-2*	hsa-miR-3198	hsa-miR-610
hsa-miR-1207-5p	hsa-miR-3125	hsa-miR-32*	hsa-miR-617
hsa-miR-1228*	hsa-miR-3131	hsa-miR-3202	hsa-miR-622
hsa-miR-1237	hsa-miR-3138	hsa-miR-320c	hsa-miR-630
hsa-miR-1254	hsa-miR-3141	hsa-miR-3614-5p	hsa-miR-642b
hsa-miR-1268	hsa-miR-3148	hsa-miR-3667-5p	hsa-miR-659
hsa-miR-1275	hsa-miR-3149	hsa-miR-3676	hsa-miR-670
hsa-miR-129-5p	hsa-miR-3161	hsa-miR-3679-3p	hsa-miR-708
hsa-miR-1305	hsa-miR-3180-3p	hsa-miR-3679-5p	hsa-miR-711
hsa-miR-1306	hsa-miR-3185	hsa-miR-373*	hsa-miR-765
hsa-miR-139-3p	hsa-miR-3187	hsa-miR-3945	hsa-miR-877
hsa-miR-151-3p	hsa-miR-3189	hsa-miR-422a	hsa-miR-892b
hsa-miR-187*	hsa-miR-3190	hsa-miR-4274	hsa-miR-595
hsa-miR-188-5p	hsa-miR-3195	hsa-miR-4294	hsa-miR-33b*
hsa-miR-18b*	hsa-miR-3197	hsa-miR-483-5p	hsa-miR-30b*
hsa-miR-1909*	hsa-miR-3198	hsa-miR-500a	
hsa-miR-202	hsa-miR-32*	hsa-miR-501-5p	
hsa-miR-206	hsa-miR-3202	hsa-miR-520b	

hsa-miR-21*	hsa-miR-320c	hsa-miR-520e	
hsa-miR-2278	hsa-miR-331-3p	hsa-miR-550a	
hsa-miR-23a*	hsa-miR-338-5p	hsa-miR-583	

We interacted the different upregulated miRNA and SNP reported in breast cancer by in-silico method. By target prediction we found 34 miRNA:SNP pairs. The obtained pairs are further filtered and finally pairs having a more significant association were screened. Siginicant pair was filtered on the basis of :

1. Number of G-U pairing present. Maximum permissible G-U pairing should be 2.
2. Maximum free energy should be > +20 and< - 20.

On the basis of these three criteria we found six significant pairs. Then the whole 3' UTR region of the finally screened genes were folded in different online RNA folding tools, ex :RNA fold, Mfold, Sfold etc., to predict the availability of the target site for the micro RNA action. There was significant difference in site availability between ancestral and mutated form of the 3' UTR region of the gene. The screened pairs were taken for the further study.

Table 2. List of SNP:miRNA pair involved in breast cancer

SNP ID and gene name	**miRNA**	**miRNA binding Site in ancestral**	**miRNA target change in mutated**	**Mutation**	**Energy variation.**
rs1259938 (MBD2)	hsa-miR-1254	8mer	Other	A/G	-27.34
rs3734805 (CCDC170)	hsa-miR-483-5p	Other	7mer-m8	A/C	-24.21
rs7441 (DNC)	hsa-miR-4835p	Other	8mer	C/T	-21

rs11662595 (HRH4)	hsa-miR 3138	Other	7mer-A1	A/G	-20
rs7539542 (ADIPOR1)	hsa-miR-3180-3p	7mer-m8	6mer	C/G	-27.35
Rs74441 (DCN)	hsa-miR-4360	Other	7mer-A1	C/T	-41.81

Conclusions:

From extensive literature studies, we have collected a total of 604 SNPs present in different genic regions which are significantly associated with breast cancer. We also collected 197 differentially expressed miRNAs in breast cancer from various databases. From the RNAHybrid analysis, we have filtered six SNP:miRNA pairs which clearly depicts the significant changes in target sites leading to alteration in miRNA targeting. For example, DCN gene coding for Decorin protein influence fibrillogenesis, inhibit the activity of TGF-beta1, inhibit cell proliferation, collagen systhesis, DNA synthesis. The presence of the SNP in DCN gene changes the miRNA targeting and a new target site is created. This creation of a new miRNA target site due to presence of SNP might be contributing to the downregulation of DCN and might be influencing Breast cancer sustcepitibility. Form the above interaction studies, we can hypothesize that occurance of SNPs in the target site of miRNA can modulate the mode of regulation of mRNA by the corresponding miRNA which ultimately might be involved in breast cancer.

References:

1. Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

2. Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

3. Gong, W.F., Zhong, J.H., Xiang, B.D., Ma, L., You, X.M., Zhang, Q.M., and Li, L.Q. (2013). Single nucleotide polymorphism 8q24 rs13281615 and risk of breast cancer: meta-analysis of more than 100,000 cases. PLoS One *8*, e60108.

4. Gray, I.C., Campbell, D.A., and Spurr, N.K. (2000). Single nucleotide polymorphisms as tools in human genetics. Hum Mol Genet *9*, 2403-2408.

5. Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674.

6. Mahdi, K.M., Nassiri, M.R., and Nasiri, K. (2013). Hereditary genes and SNPs associated with breast cancer. Asian Pac J Cancer Prev *14*, 3403-3409.

7. Mallick, B., and Ghosh, Z. (2012). Probing Evolutionary Biography of MicroRNAs and Associated Factors. Curr Genomics *13*, 144-152.

8. Nicoloso, M.S., Sun, H., Spizzo, R., Kim, H., Wickramasinghe, P., Shimizu, M., Wojcik, S.E., Ferdin, J., Kunej, T., Xiao, L., *et al.* (2010). Single-nucleotide polymorphisms inside microRNA target sites influence tumor susceptibility. Cancer Res *70*, 2789-2798.

9. Onay, V.U., Briollais, L., Knight, J.A., Shi, E., Wang, Y., Wells, S., Li, H., Rajendram, I., Andrulis, I.L., and Ozcelik, H. (2006). SNP-SNP interactions in breast cancer susceptibility. BMC Cancer *6*, 114.

10. Siomi, H., and Siomi, M.C. (2009). On the road to reading the RNA-interference code. Nature *457*, 396-404.

11. Stahlhut Espinosa, C.E., and Slack, F.J. (2006). The role of microRNAs in cancer. Yale J Biol Med *79*, 131-140.

12. Xu, Y., Gu, L., Pan, Y., Li, R., Gao, T., Song, G., Nie, Z., Chen, L., Wang, S., and He, B. (2013). Different effects of three polymorphisms in MicroRNAs on cancer risk in Asian population: evidence from published literatures. PLoS One *8*, e65123.

13. Zhang, L., Liu, Y., Song, F., Zheng, H., Hu, L., Lu, H., Liu, P., Hao, X., Zhang, W., and Chen, K. (2011). Functional SNP in the microRNA-367 binding site in the 3'UTR of the calcium channel ryanodine receptor gene 3 (RYR3) affects breast cancer risk and calcification. Proc Natl Acad Sci U S A *108*, 13653-13658.

14. Zhao, Z., Liu, C., Zeng, Y., Gu, L., Ying, M., Wang, N., Hao, B., Yao, H., Su, C., Wang, Y., *et al.* (2014). The association between the APE1 Asp148Glu polymorphism and breast cancer susceptibility: a meta-analysis based on case-control studies. Tumour Biol.