3. REVIEW OF LITERATURE
3.1. Cancer-An Introduction

Cancer – is the term derived from a Greek word "karkinos" primarily used by the father of medicine-Hippocrates, 500 BCE to describe a carcinoma resembled a crab. But, the actual history of cancer goes beyond him, where world's oldest cancer evidence hails from ancient Egypt which documented 8 cases of breast tumor occurred in 1500 BCE which also mentioned surface tumors were surgically removed as similar as today (www.cancer.gov, USA). Cancer is a harmful disease originated from aggregation of abnormal cells which are naturally immortal by escaping apoptosis through uncontrolled cell division. The cancer cells can constitute a tissue mass known as tumor. These cells are more capable of invading adjacent tissue or migrate to distant organs of the body through the blood and lymph systems generally termed as “metastasis” (Figure 3.1). According to the growth, invasiveness and metastatic potential the tumors can be further classified into two types namely (i) benign (ii) malignant. In general, the malignant tumors are recognized as more dangerous than benign since they have rapid growth and metastatic ability (American Cancer Society, USA).

![Figure 3.1. Development of Cancer](http://www.med.upenn.edu/feldserlab)

Tumorigenesis relies on heritable variation in the genetic material of somatic or germ cells. Normal cells which continuously acquire these unreplaceable genomic variations could be ultimately transformed into a cancer cell. The accumulation of the cancer cells with genomic and phenotypic changes is possible at any organ of a multicellular organism. Also, the development of cancer resemble an evolutionary process, which can originate any part of the human body and could affect any level of life. At present, more than hundred different types of cancers are known with distinct clinical and pathological features where most of them are fatal or not curable.
3.2. The Human Genome and Cancer

Decoding the human genome is a great milestone of science advancement in 21st century. The scientific venture on human genome started from the discovery of DNA sequencing by Frederick Sanger in 1977 then parallel advances in molecular cloning and high-thorough put techniques accelerated the completion of human genome project in 2001. A typical human genome gathered from an egg and sperm cells was identified to be composed of $3 \times 10^9$ base pairs (bp) of DNA (Olson et al., 1993). The genome of an individual is subjected to harbour non-heritable variations in somatic cells and heritable variations in germ cells which are responsible for the biological traits and various diseases.

![Image of cancer genome]  

**Figure 3.2. Overview on Cancer Genome**
Adopted and modified from Weinstein et al., (2013)

Historically, the carcinogenesis has been well demonstrated using chemical mutagens, biological agents like viruses, physical damages by UV rays that potentially alter the DNA sequence and can cause cancer (Cairns et al., 1981). Each cancer cell acquire an individual genome through attaining the accumulation of genomic alterations. Hence, each cancer cell may have its own genetic architecture. Apart from these variations, the genome is tightly regulated by DNA methylation, a conservatory mechanism that is also abruptly altered in cancer cells which results in misbehaving genome (Figure 3.2). The origin of a cancer genome has been pillared on three independent elements like (i) Germ line Variations (ii) Somatic Mutations (iii) DNA methylation. Each of these elements could determine the fate of genome and its stability (Stratton et al., 2009).
3.3. The Hallmarks of Cancer

The hallmarks of cancer comprise ten biological capabilities acquired during the multistep development of human tumors (Figure 3.3). Debatably the most fundamental characteristic of cancer cells includes the ability of sustain chronic proliferation. Normal tissues maintains the control of production and release of growth-promoting signals to ensuring a homeostasis of cell, whereas the cancer cells deregulates these signals. The evading signals such as mitogen signal, signals to stimulate normal cells to support tumor-associated stroma, surface receptor signals, ligand-stimulated receptor, the activation of one or another of these downstream pathways such as Ras pathway (Lemmon & Schlesinger, 2010) and constitutive activation of signalling circuits and functional impact of crosstalk between the multiple pathways radiating from growth factor receptors downstream. Defects in negative-feedback mechanisms are capable of enhancing proliferative signalling in tumor cells (Amit et al., 2007).

![Figure 3.3. Hallmarks of Cancer](image)

*Figure 3.3. Hallmarks of Cancer*
Adopted from Hanahan & Weinberg (2011)
In addition to inducing and sustaining positively acting growth-stimulatory signals, cancer cells must bypass the regulation by tumor suppressors such as TP53, RB and many others. “Contact inhibition”, a cellular phenomenon by which cells stops proliferation when cell-cell contacts formed by dense populations of normal cells propagated is established. But cancer cells evade contact inhibition and invade local or remote region or both by a process called epithelial-to-mesenchymal transition (EMT) which is associated with high grade malignancy (Korpal et al., 2008). The idea that programmed cell death by apoptosis serves as a natural barrier to cancer development is also overcome by tumor cells through triggering the levels of insulin-like growth factor 1 and 2 and by down regulating proapoptotic factors such as Bax, Bim, Puma etc. (Hanahan & Weinberg, 2011). Cancer cells require limitless replicative potential in mandate to form macroscopic tumors. Therefore, they triggers telomeres and in particular its protein subunit TERT to protect the ends of chromosomes to enable the capability for unlimited proliferation (Blasco et al., 2005).

The tumor-associated neovascularization, a process by which the development of the vasculature involves the birth of new endothelial cells and their assembly into tubes to supply nutrients and oxygen and to remove the wastes and carbon dioxide expelled by the cancer cells helps in the active development of tumor. Once angiogenesis has been activated, tumors display varied patterns of neovascularization with in (Murdoch, et al., 2008). It was strong that as carcinomas forms from epithelial tissues progressed to higher grades of malignancy with alterations in their shape, anchoring with other cells and to the extracellular matrix (ECM), revealed in local invasion and distant metastasis (Gupta et al., 2005).

Other mechanisms of tumor metastatic property may involve suppression of anti-growth signals embedded in normal tissue extracellular matrix and tumor-suppressing actions of the immune system (Kim et al., 2007). Development of genomic instability in cancer cells, which generates random mutations including chromosomal rearrangements favors the tumor formation and survival. Epigenetic mechanisms such as DNA methylation and histone modifications, clonal expansion by skipping the regulation of gene expression in both post transcriptional and post translational mechanisms and defective DNA-maintenance machinery supports the potential establishment of tumor (Kinzler & Vogelstein, 1997).
3.4. Physiology of Human Stomach

Stomach is the principal digestive organ located at upper left quadrant of the abdomen (Silen, 1967). An active stomach collects food from the esophagus and facilitates mechanical cum chemical modifications that convert complex structure of food into smaller and easily absorbable by intestines. Apart from serving a major participant in digestion, the stomach also acts as a neuroendocrine organ which produce few necessary hormones namely gastrin, ghrelin, Cholecystokinin, gastric inhibitory peptide (GIP) and pepsin. An impairment of stomach could be disastrous that eventually affects other bodily functions.

![Figure 3.4. Anatomy of a human stomach](http://www.austincc.edu)

Anatomically, the human stomach is divided into five regions: gastroesophageal (GE) junction or cardia, fundus, corpus, antrum and pylorus (Figure 3.4). The inner gastric mucosa is made up of simple columnar epithelium with numerous tubular gastric glands which are formed with four different types of cells like mucous, parietal, chief and endocrine cells. Both the fundus and corpus are rich in exocrine glands composed of hydrochloric acid-secreting mucous, parietal, and chief cells, where the antrum has been built with an alkaline-secreting surface made up of an endocrine epithelium abundant in gastrin-secreting G-cells (Soybel, 2005).
3.5. Gastric Cancer – An overview

Gastric cancer is a most frequent type of cancer that display an accountable impact on global health. In both sexes, it is the fifth most common cancer in incidence and third in mortality among all cancers worldwide (Torre et al., 2015). Stomach cancer rates are generally about twice as high in men as in women and vary widely across countries (Figure 3.5). The distribution of this morbidity varies among different geographical regions. For the past few decades, overall incidence is slightly declining in developed nations. However, the occurrence still remains a serious life threatening issue in global population. The severity of the disease could have greater influence on global population since higher incidence substantially occurs in Asia where majority of the affected individuals are diagnosed in China and other South East Asian countries. In general, incidence rates are highest in Eastern Asia (particularly in Korea, Mongolia, Japan, and China), Central and Eastern Europe, and South America and lowest in Northern America and most parts of Africa.

Figure 3.5. Global Cancer Incidence
(GLOBOCAN 2012)

In more developed countries, decreases in smoking prevalence may also account for some of the decline. Although stomach cancer is declining overall, adenocarcinoma of the gastric cardia is increasing and thought to be related to increased obesity and perhaps improvement in classification. Stomach cancer incidence is more common in people...
between the age of 60-80 years and patients below 30 years are very rare. In India, stomach cancer incidence is more prevalent among people at the age group 45-55 years in the North and 35-55 years in the South. Like the global observation, the majority of the cases are occurring in male with rates two to four times higher than females (Mallath et al., 2014). In 2010, a national wide assessment of cancer deaths in India indicated that stomach cancer is second most common fatal cancer with the mortality rate of 12.6% (Figure 3.6).

Figure 3.6. Cancer Incidence in India (A) Incidence (B) Mortality (C) Prevalence Adopted from Mallath et al., (2014)

Unlike other type of cancers, the majority of gastric cancer cases are often diagnosed at advanced stages which restrict the clinical interventions and treatment strategies. The gastric cancer development is a complex process harmonized by multiple entities including etiological factors and host genetics. The two main histologic subtypes of the disease, intestinal and diffuse type, as classified by Lauren, define two distinct entities that have different epidemiology, etiology, pathogenesis and behavior.
3.6. Etiology and risk factors

Gastric cancer development has been linked with a variety of common risk factors (Figure 3.7) namely *H. pylori* infection, Ageing, smoking, alcohol consumption and food habits (Compare *et al.*, 2010). *H. pylori*, a Gram-negative microaerophilic, spiral bacterium found in the gastric mucosa of patients with chronic atrophic gastritis. This bacteria has been identified as group I carcinogen by International Agency for Research on Cancer which is associated with an approximately two-fold increased risk of developing gastric cancer. The developing countries is suspected to be infected higher infection rate of *H. Pylori* where 50% world population is thought to be infected with this bacteria. The European Prospective Investigation into Cancer and Nutrition (EPIC) project reported that tobacco smoke has been linked to induce the development of precursor gastric lesions to intestinal metaplasia.

![Risk factors for gastric cancer](image)

**Figure 3.7. Risk factors for gastric cancer**

It has also been demonstrated that cigarette smoking could decrease prostaglandins that maintain gastric mucosal integrity. Moreover, higher incidence of *H. pylori* infection has been connected with smokers. Alcohol, a gastric irritant which is another important risk factor for gastric cancer. Consumption of alcoholic beverages could be an increased risk of stomach cancer in both men and women (Krejs, 2010). Food habits are well-known risk factors for gastric cancer where consumption of large amount of salted fish, soy sauce, pickled vegetables, cured meat and other salt-preserved foods can damage the gastric mucosa resulting in gastric carcinogenesis.
3.7. Genetics of Gastric Cancer

Gastric cancer development requires a complex genetic and environmental interactions (Figure 3.8) that contribute to its initiation and progression (McLean & El-Omar 2014). Majority of the cases are adenocarcinomas which is traditionally divided into two main subtypes namely (i) intestinal (ii) diffuse type. Both these subtype show different molecular profile and their development pathways are also distinct. Over the past few decades, advances in technology and high-throughput analysis have enabled a greater appreciation of the molecular aspects of gastric cancer pathogenesis.

![Figure 3.8. Genetics of gastric cancer](Adopted from McLean & El-Omar (2014))

**Germ line Variations- SNP**

The gastric cancer can arise by germ line or sporadic mutations. Classically the hereditary diffuse gastric cancer (HDGC) has been associated with heterozygous germ line mutation in E-Cadherin gene (CDH1). However, the familial history of gastric cancer is very low and inherited as autosomal dominant fashion. Apart from the germ line mutations, single nucleotide polymorphisms (SNPs) are known to make individual susceptible to gastric cancer. The frequency of these naturally occurring genetic variations differ between ethnic identities. A set of SNPs in the gene or genomic loci can form a haplotype which can influence the traits of the gene and affect its function. Many examples exist of polymorphic genes that increase the susceptibility to gastric cancer. Recently, advanced technologies, genome-wide association studies (GWAS) and high-throughput genetic
analysis have made information available on multiple SNPs and offering new insights into the pathogenesis of gastric cancer. The frequency of particular genetic polymorphisms, as well as which polymorphisms are most common, varies in accordance with ethnicity, and together with exposure to environmental risk factors can promote the development of malignant disease (Wadhwa et al., 2013). In gastric cancer, there are high volume of SNP data available which focused on genes involved in mucosal protection (e.g., IL1B, IL1RN, and TNF-α), carcinogen metabolism (e.g., CYP2E1 and GSTM1), DNA repair (e.g., MTHFR and XRCC1), and tumor suppressors (e.g., TP53 and CDH1). Except these genes, some other candidate genes have been widely investigated including PSCA which harbours SNPs linked with gastric cancer susceptibility.

**Somatic Variations- Mutations**

Advanced methods have revolutionized the detection of somatic variations in cancers. To date, a large spectrum of somatic alterations have been identified and reported regularly in gastric cancer where genes like *TP53*, *PIK3CA*, and *ARID1A* are identified to yield recurrent mutations (Zang et al., 2012). The accumulation of cancerous mutations could impair the biological pathway of cell adhesion.

**DNA Methylation**

DNA methylation by epigenetic modifications are more common events in pathogenic in malignant disease, including gastric cancer. This modifications could alter the expression of target genes by changing the methylation status of DNA CpG island. Epigenetic events, most obviously established by stable and heritable changes in gene expression that are not due to any alteration in the primary DNA sequence, signify the fundamental molecular principles in which genetic information is organized and read. In the past decades, it has become increasingly evident that altered epigenetic control of gene expression plays a substantial role in many different diseases, including gastric malignancies (Qu et al., 2013). Promoter methylation is an important hallmark of cancer cells, which plays a key role in the initiation and progression of tumor, including gastric cancer. Inactivation of tumor suppressor genes by promoter methylation is one of the major mechanism that ultimately leading to gastric carcinogenesis. Moreover, aberrant methylation of a number of genes has been demonstrated in gastric cancer.
3.8. Genes implicated in gastric tumorigenesis

**DNMT1**

DNA methylation plays an important role in chromatin remodelling in mammalian cells. In cancers, both hypomethylation and hypermethylation have been well documented (Jones *et al.*, 2002) where DNA methyltransferase 1 (DNMT1) overexpression have been shown to be involved in hypermethylation. In gastric cancer, both the subtypes have been identified to yield DNA hypermethylation.

**AGO4**

Argonaute 4 (AGO4) also known as EIF2C1, is a member RNA mediated gene silencing complex in mammal. In human, it is located on chromosome 1p34-p35, a genomic region is frequently lost in human cancers. The human *EIF2C1* gene is ubiquitously expressed at low to medium levels, where its expression was found to be elevated in neoplastic development.

**KLF4**

Krüppel-like factor 4 (KLF4) was one among the KLF family consists of at least 16 different members. The KLF4 was identified as zinc-finger-containing transcription factor that crucially involved in regulating embryogenesis in *Drosophila melanogaster*. Human *KLF4* (formerly known as gut-enriched KLF or epithelial zinc finger, EZF) was first identified from human umbilical vein endothelial cell (Yet *et al.*, 1998). To date, several functional domains of KLF4 have been characterized, where it has been verified to bind with promoter of a highly conserved family of genes including cytochrome P-450 drug-metabolizing enzymes. Constitutive expression of KLF4 inhibits DNA synthesis and reduces cell proliferation. KLF4 is predominantly expressed in crypt epithelium of GI tract suggesting that it is important for differentiation of gut epithelial cells and rapidly upregulated in response to vascular injury (Wei *et al.*, 2005).

**FGFR2**

Fibroblast growth factor receptor 2 belongs to FGFR receptors family. FGFR2 is associated with Fibroblast growth factor signalling and enables multiple biological activities, including fundamental developmental pathways, cellular proliferation, differentiation, motility and transforming activities. FGFR2 are expressed in normal developing cells and highly exploited in cancer cells where the deregulation of FGF
signalling is frequently observed in various solid cancers (Turner & Grose, 2010). FGFR2 amplification has been found in diffuse-type gastric cancer and reported to be involved in poor prognosis (Toyokawa et al., 2009). On the other hand, FGFR2 amplification confers hypersensitivity to FGFR inhibitor which strongly suggesting that FGFR2 amplification may be a promising molecular target for the treatment of FGFR2-amplified gastric cancer (Takeda et al., 2007).

**EGFR**

Epidermal growth factor receptor (EGFR) is a transmembrane RTK and a member of human epidermal growth factor receptors family that regulate diverse downstream signaling pathways and play an important role in development of gastric cancer. Overexpression of EGFR has been shown to be related to poor prognosis (Kim et al., 2012).

**3.9. Non coding RNA**

Eukaryotic genomes are complex and well-ordered substrates of gene transcription and RNA is not only a messenger operating between DNA and protein that was once believed. With the entry of advanced high-throughput genomic technologies such as real time quantitative PCRs, microarrays and next-generation sequencing (NGS) has resulted in an unprecedented scope and ability to detect novel transcripts, the vast majority of transcripts which seem not to be derived from annotated protein-coding genes. The foremost RNA partakers in gene expression, the ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs), were discovered in the 1950s and their chunk in protein synthesis firmly established (Cech, 2014). It was not demonstrated the existence of non-coding RNAs (ncRNAs) until the discovery of small nuclear RNAs (snRNAs) as a possible players in the excision of introns from the immature mRNAs (Peebles, 1986) and later small nucleolar RNAs (snoRNAs) joined the clamour.

In the early 2000s the non-coding RNA science jumped to a revolutionary high place with the discovery of microRNAs or miRNAs emphasising the prominence of posttranscriptional events in gene expression particularly in eukaryotic organisms. There are more than 2000 different miRNAs which have been discovered and estimated in human genome. miRNAs play a critical role in cell development (Esteller, 2011). Also it is shown to be involved not only in tumor promotion but also as tumor suppression functions. Since then, several microRNAs have developed as candidate component of oncogenes and tumor
suppressor networks. In cancer, generally up-regulation of oncogenic miRNAs results in down-regulation of tumor suppressor genes while down-regulation of tumor suppressive miRNAs results in up-regulation of oncogenes.

In recent days, deep sequencing has revealed the being of thousands of long non-coding RNAs (lncRNAs) with a variety of functions in both gene expression and remodelling of the eukaryotic genome (Cech, 2014). Thought there are frequent discoveries of new functions of “classic” ncRNAs such as ribosomal RNAs (rRNAs), ribozymes, transfer RNAs (tRNAs), snRNAs, snoRNAs, eRNAs and telomere-associated RNAs well look over in literatures, the recent interest in lncRNAs is deeply rooted in biology’s longstanding concern with the evolution and function of genomes (Figure 3.8).

According to LNCipedia 2.0, the latest version of this long noncoding RNA database, there are at present 32,183 annotated lncRNAs in human (LNCipedia 2.0, 2014). Based on the length of the non-coding RNAs they are classified into two major categories, long non-coding RNAs which are >200 nucleotides in length and miRNAs/endogenous small interfering RNAs (endo-siRNAs) are covered under small ncRNAs which are below 200 nucleotides in length (Clark et al., 2012). Recent reports have put forward that lncRNAs by executing as signals, decoys, guides and scaffolds and act as a repressor or activator to modulate the process of gene transcription and translation thereby it regulates gene expression, whereas the miRNAs regulates gene expression only by post transcriptional level (Wang et al., 2011). In many cases, miRNAs negatively regulates target expression at the translational level by bind to the target 3’ UTRs through imperfect complementarity at multiple sites (Bartel, 2004).

In contrast, long non-coding RNAs acts as a decoy molecule that sequesters miRNAs to revoke their functions in regulating target gene expression. Poliseno et al., (2010) reported that the 3’ UTR of PTENP1, a pseudogene transcript which binds to the same set of regulatory miRNA sequences that target the tumor suppressor PTEN mRNA. As a result, the miRNA inhibition of PTEN expression is attenuated and PTEN levels are increased. There are very limited evidences available on the miRNA transcriptional regulation of the lncRNA transcription.
Figure 3.9. Proportion of Non-coding RNAs. The pie chart represents the major category of non-coding RNAs which includes miRNAs and IncRNAs. Adopted from Sacco et al., (2011).

Figure 3.10. Composition and quantities of IncRNA. Pie chart of the composition and quantities of IncRNA, transcript of unknown coding potential (TUCP), expressed pseudogene, read-through and protein-coding genes in the MiTranscriptome assembly. Pie charts of the number of IncRNA and TUCP genes that are unannotated versus annotated relative to reference catalogues (top right) and intragenic versus intergenic (bottom right) Adopted from Iyer et al., (2015).
3.10. Long non-coding RNA

Biogenesis of Long non-coding RNAs

The large proportion eukaryotic genome is transcribed into huge array of RNA molecules differing in size, abundance and protein-coding and non-coding capability. Only ~1% of the human genome codes for proteins, and ~4%–9% that is transcribed but yet their functions are greatly unknown (Mercer et al., 2009). The existence of individual lncRNAs such as H19 and Xist has been known since the 1980s (Brannan, 1990. & Brockdorff, 1992). Almost all lncRNA species are >200 bp in length and these include antisense RNAs, intronic RNAs and intergenic lncRNAs (Faulkner, 2009). Nearly 40% (3934 lncRNA genes, 5361 transcripts) of GENCODE lncRNAs overlaps protein-coding gene loci (Derrien, 2012). The majority of lncRNAs are transcribed by RNA polymerase II and also they poses epigenetic signatures common to protein-coding genes like 5’ caps, histone modifications associated with Pol II transcriptional elongation such as trimethylation of histone 3 lysine 4, and polyadenylation signals alike mRNAs (Guttman et al., 2009). Like the transcription of coding RNAs, transcription of lncRNAs also occurs from an independent gene promoter and therefore it is not coupled to the transcription of a nearby or associated parental gene (Gibb et al., 2011). LncRNA expression is controlled by both transcriptional and epigenetic factors, which is similar to the protein coding genes except the epigenetic regulation at the level of DNA methylation is evidently dissimilar (Sati et al., 2012). Precisely, lncRNAs demonstrate a comparatively lower level of expression but much more tissue-specific than protein-coding genes, signifying illustrious and regulatory role of lncRNAs (Derrien et al., 2012).

Figure 3.11. Classification of lncRNAs based on genomic position relative to the nearest protein-coding gene. Long intergenic non-coding RNA (lincRNA) genes do not overlap neighbour protein coding genes. Antisense lncRNAs are transcribed from the strand opposite strand.
Figure 3.12. Various mechanisms of lncRNA function at epigenetic, transcriptional and post-transcriptional levels. Adopted from Yang et al., (2014).
A eukaryotic genome is capable of producing numerous amount of non-protein-coding RNA species that show complex overlapping patterns of expression and regulation. Although only a tiny amount of lncRNAs have been functionally well characterized to date, they have been shown to control every level of the gene expression program by implicating in posttranscriptional level through controlling processes like protein synthesis, RNA maturation, and transport and in transcriptional gene silencing through regulating the chromatin structure (Whitehead et al., 2009).

**Epigenetic Regulation**

Regulation of IGF2 by lncRNAs through highly specific to a particular gene (Matouk et al., 2007) or, it can perform on a wide chromosomal region like X-chromosome inactivation in females by XIST, perhaps the most well studied lncRNAs which is transcribed from the inactivated X chromosome (Brown, 1991 & Lee 2009). XIST propagate epigenetic silencing of an individual X chromosome through its double-hairpin RNA motif in the RepA domain at first exon by binding to PCR2 (Zhao et al., 2008).

**Chromatin modifications**

Long non-coding RNAs contribute in targeted gene silencing through chromatin remodelling. Recent reports identify ANRIL (CDKN2B anti sense) as another candidate lncRNA which interacts with more than one chromatin remodelling complex to induce silencing of target genes. *In vitro* data have suggested that ANRIL an antisense RNAs transcribed from the CDKN2A and CDKN2B functions to repress the INK4A/INK4B isoforms, but not ARF which is mediated through direct binding to CBX7 applying repressive histone modifications to the locus (Yap et al., 2010).

**Genomic Imprinting**

Numerous *cis*-regulatory lncRNAs, including H19, AIR, COLDAIR, HOTTIP and KCNQ1OT1, are also functionally related through their participation in epigenetic imprinting regions. In humans, maternally and paternally expressed lncRNAs H19 and KCNQ1OT1 respectively, retain silencing of the IGF2 and KCNQ1 genes on those alleles (Weksberg et al., 2001; Bliek et al., 2001). The functional data on H19 shows it has been linked to both oncogenic by direct activation by c-Myc and tumor suppressive qualities through downregulation by p53 and during prolonged cell proliferation (Gabory et al., 2010; Pantoja et al., 2005).
Fig 3.13. LncRNA mediated genomic imprinting. Model showing the cis-regulation of gene expression results in local control of genes neighbouring, or on the same chromosome as, lncRNA transcription. H19 and KCNQ1OT1 are imprinted lncRNAs on chromosome 11 associated with allele-specific expression of IGF2 and KCNQ1.

**Enhancing/repressing gene promoters**

In addition to aiding epigenetic changes that influence gene transcription, some lncRNAs contribute to gene regulation by persuading the activity of gene enhancers. MEG3, a maternally-expressed imprinted lncRNA on Chr14q32, has been shown to activate p53 gene expression (also known as p53 co-activator lncRNA) and facilitate p53 signalling, including enhancing p53 binding to target gene promoters (Zhou et al., 2007). LncRNA transcripts can activate transcription factors via allosteric interactions, such as the activation of the Dlx5/6 enhancer by the lncRNA Evf2 in trans (Berghoff et al., 2013).

**Guides and scaffolds**

When lncRNAs bind to proteins and direct the localization of RBP complex to specific target for chromatin modifications, thus they are working as “molecular guides”. This mode of deed is quite complicated. lncRNAs may interact with several effector molecules such as the trithorax group proteins (TxG), the polycomb group proteins (PcG), and the common set of transcription factors for chromatin modifications (Wang & Chang 2011). HOTAIR was found to regulate HoxD cluster genes in a trans-regulatory mechanism. HOTAIR is primarily associated with polycomb repressive complex 2 (PRC2),
where it acts along with trimethylate H3K27 to repress transcription of specific genes (Rinn et al., 2007; Gupta et al., 2010). Ge et al., (2013) reported that HOTAIR increases the H3K27 methylation in the WIF-1 promoter and induces its silence apart from the hypermethylation of the gene promoter clearly shows that the lncRNAs are involved in epigenetic mechanism of gene regulation.


In breast cancer, HOTAIR overexpression facilitates increasing PRC2 level and its recruitment to the genomic positions of target genes which in turn mediates the epigenetic repression of PRC2 target genes through changes in chromatin structure. The characterization of HOTAIR brought widespread attention to trans-regulatory lncRNAs into prospect (Rinn et al., 2007).

**LncRNAs as a Decoy**

GAS5 induces apoptosis and suppresses cell proliferation and in human breast tumors its expression is downregulated (Mourtada-Maarabouni et al., 2009). GAS5 modulates cell survival and metabolism by antagonizing the glucocorticoid receptor (GR) by interacting with the GR DNA-binding domain (DBD) and represses GR-induced genes thereby serving as a decoy that prevents GR binding to target DNA sequences (Kino et al., 2010). PANDA which is expressed only in p53-positive cells may also act as a decoy by interacting with the transcription factor NF-YA down-regulating the expression of pro-apoptotic genes and enabling cell-cycle arrest (Sacco et al., 2012). During DNA damage CCND1, a long ncRNAs associated with the cyclin D1 gene promoter can bind to RNA binding protein TLS (translocated in liposarcoma) and modulate the activities by an allosteric function. TLS later inhibits the histone acetyltransferase activities of CREB binding protein and p300 to silence cyclin D1 expression (Wang et al., 2010).
**RNA editing**

Post-transcriptional processing of mRNAs is also a dynamic make up to gene expression, while many lncRNAs function by regulating gene transcription. Nuclear paraspeckle, a sub-cellular protein granules compartment found in the interchromatin space serve as storage sites for mRNA earlier to its pass on to the cytoplasm for translation. CTN-RNA a polyadenylated nuclear ncRNA, counterpart to the protein-coding murine CAT2 (mCAT2) gene under stress conditions cleavage of CTN-RNA non-coding region to convert itself into mCAT2 coding transcript resulted in increased mCAT2 protein. MALAT1 and NEAT1 are overexpressed in cancer cells are involved in mRNA splicing and nuclear paraspeckle function (Prensner & Chinnaiyan 2011). Interaction between MALAT1 and the serine/arginine (SR) splicing factors effects the distribution and phosphorylation of SR, subsequently leads to changes of the alternative splicing of a set of endogenous pre-mRNAs (Tripathi et al., 2010). LncRNAs can also regulate mRNA alternative splicing. For instance, ZEB2-AS transcribed from the antisense orientation to ZEB2 can bind the ZEB2 pre-mRNA, avoiding the splicing of a 5′ UTR IRES-containing intron, thus increases protein levels of ZEB2 (Beltran et al., 2008).

**miRNA sponging**

Recent work on mechanisms of RNA regulation by RNA-RNA interactions between ncRNAs and mRNA sequences such as miRNA regulation of mRNAs, as sequence homology between the ncRNA and the mRNA is important to the regulatory process. The tumor suppressor gene phosphatase and tensin homolog (PTEN) which is a negative regulator of the PI3K-Akt pathway was epigenetically silenced in several cancers. The severity of epithelial cancers depends upon the dosage of PTEN expression, indicating that PTEN gene is critical for maintaining cellular homeostasis. PTEN expression has been found to be post transcriptionally regulated by a PTEN pseudogene. PTENP1 which was found to sequester several PTEN-targeting miRNAs by acting as a miRNA sponge thereby increased PTEN mRNA stability and increased amounts of PTEN protein (Johnsson et al., 2013). Similar mechanism is seen in case of STAU1, a RNA degradation protein, binds to protein-coding mRNAs that interact with IncRNAs containing ancestral Alu repeats forming an IncRNAs and mRNAs partially hybridize. This IncRNA-mRNA complimentary binding forming double-stranded RNA complexes that then recruit STAU1 to implement RNA degradation (Gong & Maquat 2011).
Figure 3.1. Mechanisms of lncRNA function. a. Gene expression regulation may occur through direct lncRNA-mRNA interactions which arise from hybridization of homologous sequences and can serve as a SIGNALLING for STAU1-mediated degradation of the mRNA. b. RNA molecules, including mRNAs, pseudogenes, and ncRNAs, can serve as molecular sponges for miRNAs. Adopted from Prensner et al., (2011).

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>lncRNA:chromatin regulators</td>
<td>SRA/CTCF</td>
<td>Enhances insulator function of CTCF</td>
</tr>
<tr>
<td></td>
<td>HOTAIR, lSD1-CoREST</td>
<td>Targets the LSD1 complex to demethylate H3K4me2</td>
</tr>
<tr>
<td></td>
<td>Xist, PRC2</td>
<td>Targets PRC2 either in cis or trans to mediate H3K27 methylation</td>
</tr>
<tr>
<td>lncRNA:TIs</td>
<td>Gas5/Glucocorticoid receptor</td>
<td>Titrates GR away from target genes</td>
</tr>
<tr>
<td></td>
<td>PANDA: NF-YA</td>
<td>Titrates NF-YA away from apoptotic genes</td>
</tr>
<tr>
<td>lncRNA (mRNA)</td>
<td>H19 (mIR-675)</td>
<td>Produces miR-675</td>
</tr>
<tr>
<td></td>
<td>lncRNA-202 (miR-31)</td>
<td>Produces miR-31</td>
</tr>
<tr>
<td></td>
<td>H19:cmiR-372</td>
<td>Competes for binding miR-372 with protein coding counterparts</td>
</tr>
<tr>
<td></td>
<td>CDH1:cmiR-671</td>
<td>Degrades CDH1-AS</td>
</tr>
<tr>
<td></td>
<td>PTEN:miR20/19</td>
<td>Competes with PTEN for binding with miRNAs</td>
</tr>
<tr>
<td>lncRNA:RNA</td>
<td>NAT-ZEB1</td>
<td>Alternative splicing</td>
</tr>
<tr>
<td></td>
<td>LinRNA-p21:JUNB</td>
<td>Induces degradation of target mRNA</td>
</tr>
<tr>
<td></td>
<td>1/2:3-isoRNA:2MD mRNA</td>
<td>Induces degradation of target mRNA</td>
</tr>
<tr>
<td></td>
<td>LinRNA-p21:JUNB</td>
<td>Induces degradation of target mRNA</td>
</tr>
<tr>
<td></td>
<td>LinRNA-p31:HuR</td>
<td>Induces degradation of LinRNA-p31</td>
</tr>
<tr>
<td></td>
<td>ncRNA:CDX1-1/5S</td>
<td>Allotinically binds TLU and changes its activity</td>
</tr>
<tr>
<td></td>
<td>MALAT1:serine/arginine (SR) protein</td>
<td>Modulates SR splicing factor phosphorylation and thus the downstream target splicing</td>
</tr>
</tbody>
</table>

Table 3.1. Examples of how lncRNA works as a crucial element and platform for complicated interactions. Adopted from Yang et al., (2014).

3.11. Long Non-coding RNAs and Cancer

In a molecular perspective, cancer is a genetic disease due to aberrant expression and function of tumor suppressor and oncogenic genes. LncRNAs have aroused as a crucial regulator in more or less all aspects of biology and accumulating evidence put forward that lncRNAs play significant role in tumorigenesis adding a new layer of complexity to the molecular architecture of human cancers. Genetic, epigenetic and transcriptional regulatory mechanisms have been clarified to involve in lncRNA deregulation in cancers.
Chromosomal translocations, copy-number alterations, nucleotide expansions, and single nucleotide polymorphisms (SNPs) represents molecular disorders acquired by cancer cells during neoplastic transformation. In addition loss of heterozygosity (LOH) of the maternal or paternal allele is also seen in cancers.

Differential expression of lncRNAs is increasingly documented as a hallmark feature in cancer. Many recent studies have identified the role of lncRNAs which are involved in p53 regulation or have a role in carcinogenesis or tumor growth. There are very few direct studies pointing the role of non-protein coding RNAs association with the Gastric cancer development are available till date. Although the roles of lncRNAs in gastric cancer have just begun to be revealed, its potentiality as a tumor inducer or suppressor will bring a board outlook for the molecular pathogenesis of gastric cancer.

**Figure 3.16. Dysregulation of lncRNAs in cancer cells and alteration of gene expression**
Adopted Yang et al., (2014)

After a circulating tumor cell has survived the vasculature and adhered to the endothelium of a target tissue, it must cross the endothelial barrier in order to colonize that distal tissue, the process called metastasis. With the proposed hallmarks, cancer obtains the capacities of initiation, growth and metastasis. The network leading cancer metastasis is far from conclusion. Deregulation of these genes and how they confer the cancer cells immortal capacity are still below par understood. Moving our understanding on cancer much deeper; with accumulating evidences, the involvement of lncRNAs in the process of tumor development and metastatic induction is clear.
Table 3.2. List of long non-coding RNAs with their function, associated cancer types and mechanism of action. Adopted from Prensner (2011).

<table>
<thead>
<tr>
<th>IncRNA</th>
<th>Function</th>
<th>Cancer Type</th>
<th>Cancer Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>HULC</td>
<td>Biomarker</td>
<td>Hepatocellular</td>
<td>Not known</td>
</tr>
<tr>
<td>PCA3</td>
<td>Biomarker</td>
<td>Prostate</td>
<td>Not known</td>
</tr>
<tr>
<td>ANRIL/p53AS</td>
<td>Oncogenic</td>
<td>Prostate, Leukemia</td>
<td>Suppression of senescence via INK4A</td>
</tr>
<tr>
<td>HOTAIR</td>
<td>Oncogenic</td>
<td>Breast, hepatocellular</td>
<td>Promotes metastasis</td>
</tr>
<tr>
<td>MALAT1/NEAT2</td>
<td>Oncogenic</td>
<td>Lung, prostate, breast, colon</td>
<td>Unclear</td>
</tr>
<tr>
<td>PCAT-1</td>
<td>Oncogenic</td>
<td>Prostate</td>
<td>Promotes cell proliferation; inhibits BRCA2</td>
</tr>
<tr>
<td>PCGEM1</td>
<td>Oncogenic</td>
<td>Prostate</td>
<td>Inhibits apoptosis; promotes cell proliferation</td>
</tr>
<tr>
<td>TUC338</td>
<td>Oncogenic</td>
<td>Hepatocellular</td>
<td>Promotes cell proliferation and colony formation</td>
</tr>
<tr>
<td>uc.73a</td>
<td>Oncogenic</td>
<td>Leukemia</td>
<td>Inhibits apoptosis; promotes cell proliferation</td>
</tr>
<tr>
<td>H19</td>
<td>Oncogenic; Tumor suppressive</td>
<td>Breast, hepatocellular</td>
<td>Promotes cell growth and proliferation; activated by cMYC; downregulated by prolonged cell proliferation</td>
</tr>
<tr>
<td>GAS5</td>
<td>Tumor suppressive</td>
<td>Breast</td>
<td>Induces apoptosis and growth arrest; Prevents GR-induced gene expression</td>
</tr>
<tr>
<td>linc-q21</td>
<td>Tumor suppressive</td>
<td>Mouse models of lung, sarcoma, lymphoma</td>
<td>Mediates p53 signaling; induces apoptosis</td>
</tr>
<tr>
<td>MEG3</td>
<td>Tumor suppressive</td>
<td>Meningioma, hepatocellular, leukemia, pituitary tumors</td>
<td>Mediates p53 signaling; inhibits cell proliferation</td>
</tr>
<tr>
<td>PTENP1</td>
<td>Tumor suppressive</td>
<td>Prostate, colon</td>
<td>Binds PTEN-suppressing miRNAs</td>
</tr>
</tbody>
</table>

Abbreviations: Polycomb Repressive Complex 1, PRC1; Polycomb Repressive Complex 2, PRC2; Glucocorticoid Receptor, GR

Table 3.3. Functions of selected IncRNAs highlighted in various research studies.

<table>
<thead>
<tr>
<th>S. No.</th>
<th>IncRNA</th>
<th>Functions</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NEAT1</td>
<td>Transcriptional regulator, Tumor Progression</td>
<td>Zeng C et al., 2014</td>
</tr>
<tr>
<td>2</td>
<td>UCA</td>
<td>Cell proliferation, metastasis</td>
<td>Zheng Q et al., 2015</td>
</tr>
<tr>
<td>3</td>
<td>PANDA</td>
<td>Apoptosis on DNA damage</td>
<td>Hung T et al., 2011</td>
</tr>
<tr>
<td>4</td>
<td>AP5M1</td>
<td>Cell death</td>
<td>Lee MR et al., 2008</td>
</tr>
<tr>
<td>5</td>
<td>LINC312</td>
<td>Lymph node metastasis</td>
<td>Zhang W et al., 2013</td>
</tr>
<tr>
<td>6</td>
<td>FALEC</td>
<td>BMI1 associated, negative regulator of p21</td>
<td>Hu X et al., 2014</td>
</tr>
<tr>
<td>7</td>
<td>LINCRO</td>
<td>miR145 CeRNA for P53, hypoxia induced, reprogramming</td>
<td>Zhang A et al., 2013</td>
</tr>
<tr>
<td>8</td>
<td>PTENP1-AS</td>
<td>CeRNA PTEN</td>
<td>Johnsonson P et al., 2013</td>
</tr>
<tr>
<td>9</td>
<td>ZEB2-AS1</td>
<td>Anti-sense ZEB1</td>
<td>Strausberg RL et al., 2002</td>
</tr>
<tr>
<td>10</td>
<td>HOTAIR</td>
<td>EMT and Invasion</td>
<td>Rinn JL et al., 2007</td>
</tr>
<tr>
<td>11</td>
<td>MALAT1</td>
<td>Tumor metastasis</td>
<td>Kryger R et al., 2012</td>
</tr>
<tr>
<td>12</td>
<td>CCAT1</td>
<td>Metastasis of tumor cell</td>
<td>Kam Y et al., 2014</td>
</tr>
<tr>
<td>13</td>
<td>ANRIL</td>
<td>DNA damage response, ATM-E2F1</td>
<td>Wan G et al., 2013</td>
</tr>
<tr>
<td>14</td>
<td>H19</td>
<td>Cell proliferation</td>
<td>Brannan CI et al., 1990</td>
</tr>
<tr>
<td>15</td>
<td>MEG3</td>
<td>Imprinting</td>
<td>Qin R et al., 2013</td>
</tr>
<tr>
<td>16</td>
<td>NBAT1</td>
<td>Inhibitor of cell proliferation and differentiation</td>
<td>Pandey GK et al., 2014</td>
</tr>
<tr>
<td>17</td>
<td>POU3F3</td>
<td>Methylation</td>
<td>Guo H et al., 2014</td>
</tr>
<tr>
<td>18</td>
<td>GAS5</td>
<td>Growth arrest</td>
<td>Cao S et al., 2014</td>
</tr>
<tr>
<td>19</td>
<td>TUG1</td>
<td>Proliferation signal</td>
<td>Han Y et al., 2013</td>
</tr>
<tr>
<td>20</td>
<td>7SK</td>
<td>snRNA involved in transcription</td>
<td>Peterlin BM et al., 2012</td>
</tr>
</tbody>
</table>
H19

It is expressed from the maternal allele and has a pivotal role in genomic imprinting during cell growth and development. The abnormal expression of H19 could cause loss of imprinting in many cancers. This lncRNA has been linked to oncogenic and tumor suppressor properties. The c-Myc induces the expression of H19 in different cell types where H19 potentiates tumorigenesis. In addition c-Myc also down-regulates expression of IGF2 imprinted gene. H19 transcripts are precursors for miR-675 which functionally down-regulates the tumor suppressor gene for retinoblastoma in human colorectal cancer. Supporting a dynamic role for H19 in metastasis, both silencing and over-expression of H19 was shown to modulate metastatic behaviour in bladder cancer and was significantly higher expressed in liver metastases compared to the primary tumors (Fellig et al., 2005; Luo et al., 2013). The H19 explanation for its high expression in human metastases promoter is known to be negatively regulated by p53 and positively regulated by c-Myc, E2F and HIF1-α (Berteaux et al., 2005; Amit et al., 2010).

GAS5

Growth Arrest-Specific 5 is widely expressed in embryonic and adult tissues. Expression is almost undetectable in growing leukemia cells and abundant in saturation density-arrested cells. GAS5 functions as a starvation or growth arrest-linked riborepressor for the glucocorticoid receptors by binding to their DNA binding domain inhibiting the association of these receptors with their DNA recognition sequence. This suppresses the induction of several responsive genes including the gene encoding cellular inhibitor of apoptosis 2 (cIAP2), reducing cell metabolism and synthesizes cells to apoptosis. GAS5 has been shown to induce apoptosis directly or indirectly in the prostate and breast cancer cell.

MEG3

LncRNA MEG3 is a transcript of the maternally imprinted gene. In normal pituitary cells MEG3 is expressed, the loss of expression is observed in pituitary adenomas and the majority of meningiomas and meningioma cell lines. MEG3 activates regulation of tumor suppressor protein p53 (Zhou et al., 2007). Normally, p53 protein levels are extremely low due to its rapid degradation via the ubiquitin-proteasome pathway. The ubiquitination of p53 is mainly mediated by MDM2, an E3 ubiquitin ligase. MEG3 down-regulates MDM2 expression, which suggests that MDM2 down-regulation is one of the mechanisms whereby
MEG3 activates p53. MEG3 significantly increases p53 protein level and stimulates p53-dependent transcription. MEG3 enhances p53 binding to target promoters such as GDF15 but not p21 and is also able to inhibit cell proliferation in the absence p53, suggesting that MEG3 is a p53 dependent and independent tumor suppressor.

**PTENP1-AS**

PTENP1 has been demonstrated to function as a tumor suppressor in several cancer cells. However, its expression and biological roles in gastric cancer (GC) have not yet been investigated. It has been shown to be partly associated with DNA hypermethylation, and lower PTENP1 expression was associated more advanced stage deeper invasion depth and lymphatic metastasis which suggested that PTENP1 could regulate gastric cancer cell proliferation (Guo et al., 2015). Loss of PTEN function leads to phosphorylation of PI3K and activation of the Akt pathway, subsequently initiates cancer cell invasion and migration. PTENP1 antisense (PTENP1-AS) transcript inhibits PTENP1 expression through cis (RNA-RNA binding by CeRNA mechanism) and trans (PRC2-dependent silencing) mechanisms, thus playing an oncogenic role (Tang et al., 2006; Tay et al., 2011).

**TUG1**

Taurine-upregulated gene 1 (TUG1), a 7.1-kb lncRNA, recruiting and binding to polycomb repressive complex 2 (PRC2), is generally downregulated in non-small cell lung carcinoma (NSCLC) tissues. Experiments revealed that TUG1 expression was induced by p53, and luciferase and chromatin immunoprecipitation (ChIP) assays confirmed that TUG1 was a direct transcriptional target of p53. TUG1 knockdown significantly promoted the proliferation in vitro and in vivo. Moreover, the lncRNA-mediated regulation of the expression of HOX genes in tumorigenesis and development has been recently receiving increased attention. Interestingly, inhibition of TUG1 could upregulate homeobox. Recent evidence highlights long noncoding RNAs (lncRNAs) as crucial regulators of cancer biology that contribute to tumorigenesis. LncRNA TUG1 was initially detected in a genomic screen for genes upregulated in response to taurine treatment in developing mouse retinal cells. TUG1 could affect cell proliferation through epigenetically regulating HOXB7. Increased TUG1 is correlated with outcomes in gastric cancer where the knockdown experiments revealed that knockdown of TUG1 repressed the proliferation.
both in vitro and in vivo which may serve as a candidate prognostic biomarker and target for human gastric cancer (Zhang et al., 2016).

**PANDA**

It serves as biomarker and involves in development of multiple cancers. Till date, role of PANDAR in gastric cancer is still unknown. The expression of PANDAR was significantly increased in gastric cancer tissues which was correlated with depth of invasion, TNM stage and lymphatic metastasis. Importantly, high expression of PANDAR could serve as an independent unfavorable prognostic role in gastric cancer (Ma et al., 2016). One such examples of tumor initiation is by PANDA. PANDA (~1.5-kb transcript) is located ~5 kb upstream of the CDKN1A (p21) transcription start site, is evolutionarily conserved which is specifically induced by DNA damage in a p53-dependent manner and mediates anti-apoptotic functions. Promoter regions of p53-dependent cell death genes are categorized by the presence of NF-YA binding site, which interacts with lncRNA PANDA in a highly specific manner thereby PANDA promotes cell survival by interfering with the apoptotic gene expression program (Morachis et al., 2010). Malfunction of p53 leads to over activation of PANDA and thereby the cancer cells inhibits apoptotic gene and attains immortal stage.

**MALAT1**

*MALAT1* is an lncRNA that is deregulated in many human cancers and shown to be retained specifically in the nucleus in nuclear speckles (Hutchinson et al., 2007), domains that are thought to be involved in the assembly, modification, and/or storage of the pre-mRNA processing machinery. Studies of MALAT1 expression in different cancers have linked with shorter metastasis-free survival (MFS), deeper tissue invasion, higher histological grade and shorter overall survival (OS) (Crea et al., 2014). About 30% of breast neoplasms shows over-expression of HOTAIR suggestively predicts shorter MFS and OS independently of tumor size, stage, and hormone receptor status and therefore, positively correlations between HOTAIR expression and lymph node metastasis (Gupta et al., 2010). A recent study showed that CCAT2 increases cell invasion and motility *in vitro* and *in vivo*, correlating with shorter MFS in colon and breast cancer (Ling et al., 2013). A notorious example of such an oncogenic lncRNA is metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) which is found to be over-expressed in lung
Review

cancer metastases (Ji et al., 2003). MALAT1 has been directly linked to cancer metastasis and alternative splicing (Beltran et al., 2008; Tripathi et al., 2010).

**CCAT1**

Colon cancer-associated transcript 1 (CCAT1) is a 2628 nucleotide-lncRNA and located in the vicinity of a well-known transcription factor c-Myc. CCAT1 has been found to be upregulated in many cancers, including gastric carcinoma and colonic adenoma-carcinoma. Moreover, c-Myc could promote CCAT1 transcription by directly binding to its promoter region, and upregulation of CCAT1 expression in colon cancer cells promoted cell proliferation and invasion. These data suggest that c-Myc-activated lncRNA CCAT1 expression contribute to tumorigenesis and the metastatic process (Yang et al., 2013).

**HOTAIR**

HOTAIR is a known long non-coding RNA which has recently been associated with the progression of some cancer types. It has been reported that HOTAIR expression is correlated with SUZ12 expression level and therefore may affect the epigenetic state of cancer tissues. Aberrant expression of HOTAIR was associated with TNM staging and lymph node metastasis of gastric tumors suggest HOTAIR may be involved in gastric cancer progression (Hajjari et al., 2013). Expression of HOTAIR are observed higher in cancerous tissues than in adjacent normal mucosa which highlight that HOTAIR expression may serve as a potentially important disease biomarker for the identification of high-risk gastric cancer patients (Zhang et al., 2015). Chromatin-associated lncRNA, HOTAIR, has attracted attention due to its profound pro-metastatic properties through targeting both PRC2 and LSD1/CO-REST repressive complexes to anti-metastatic loci and persuades an epigenetic reprogramming augmenting cellular plasticity (Gupta et al., 2010).

**LINCOR**

LINCOR (regulator of reprogramming) is located at 18q21a and modulator of ESCs maintenance and hypoxia-signaling pathways in hepatocellular cancer cells. LINCOR expression levels were identified to be decreased in all somatic cancer cell lines compared to stem cells or cells with stem cell-like capabilities, but the cancerous tissues have displayed increased LINCOR expression compared to corresponding normal tissues. (Rezaei et al., 2015). RNA expression profiling identified that LINCOR is a hypoxia-
responsive lncRNA which is increased in cancer cells that is functionally linked to hypoxia signaling mechanistic role for the extracellular transfer of linc-RoR in intercellular signalling to promote cell survival during hypoxic stress (Takahashi et al., 2014).

**LncRNAs are regulated by p53**

The p53 act as master controller of the lncRNA expression in both normal developing cells and cancer cell. H19 has been lately studied and found to be upregulated in many tumors and reported as negative regulator of p53. Also, it has been found that ectopic expression of H19 increased cell proliferation and its promoter is efficiently repressed by p53. Upregulation of the lncRNA H19 contributes to tumorigenesis through p53 activity regulation (Dugimont et al., 1998). MEG3 a maternally expressed lncRNA during development with higher levels in the paraxial mesoderm induces an accumulation of p53 protein. Stimulating the transcription from a p53-dependent promoter by MEG3 and thereby selectively regulates the expression of p53 targeted downstream genes (Zhou et al., 2007).

![Pathway showing the association of lncRNAs with p53 which determines tumorigenesis regulation](image)

Fig 3.17. Pathway showing the association of lncRNAs with p53 which determines tumorigenesis regulation. Adopted from Baldassarre & Masotti (2012).

Gene deletion, promoter hypermethylation and hypermethylation of other intergenic regions accounts for MEG3 lncRNA loss in cancers. Silencing or over-expression of lincRNA-RoR correlated directly with iPSC colony formation. LincRNA-RoR negatively regulates p53 expression, thereby mediating cell cycle arrest and apoptosis (Loewer et al., 2010). LincRNA-RoR interacts with phosphorylated heterogeneous nuclear...
ribonucleoprotein I (p-hnRNPI) to facilitate p53 repression through translational regulation in the perspective of DNA damage (Zhang et al., 2013).

TUG1 was a direct transcriptional target of p53 through interaction with the putative p53 response element in the promoter region of TUG1. Moreover, TUG1 could regulate cell growth both in vitro and in vivo (Zhang et al., 2014).

3.13. MicroRNA

MicroRNAs (miRNAs) are a family of 21–25-nucleotide small RNAs that binds to the target mRNAs through their seed sequence match and negatively regulate gene expression at the post-transcriptional level (Ambros, 2003). MiRNAs are shaped as a separate species from a specific precursor RNAs encoded in the genome. The structure of the primary miRNA transcript and the acknowledgement of this precursor by a nuclear processing machinery governs the structure and sequence of mature miRNAs (Lai, 2003). Around 50% of mammalian miRNA loci are found in close proximity to other miRNAs. These clustered miRNAs are transcribed from a single polycistronic transcription unit (TU), while some may be remarkable cases in which individual miRNAs are derived from separate gene promoters (Lagos-Quintana et al., 2003). Few miRNAs are generated from non-coding TUs, remaining are encoded in protein-coding TUs. Approximately 40% of miRNA loci are located in the intronic region of non-coding transcripts, whereas ~10% are placed in the exonic region of non-coding TUs (Pulito et al., 2014).

The transcription of most miRNA genes is mediated by RNA polymerase II (Pol II), although a minor group of miRNAs that are associated with Alu repeats can be transcribed by RNA Pol III (Canella et al., 2010). The following the transcription of primary miRNA by RNA Pol II, cleavages of miRNA maturation are catalysed by two RNase-III enzymes, Drosha and Dicer (Lee et al., 2003; Krol et al., 2010). Both are dsRNA-specific endonucleases that generate 2-nucleotide-long 3′overhangs at the cleavage site. Predominantly localized in the nucleus, Drosha contains two tandem RNase-III domains, a dsRNA binding domain and an amino-terminal fragment of unidentified function. Regardless of the diverse primary sequences of pri-miRNAs, Drosha cleaves them into ~70-bp pre-miRNAs that consist of an imperfect stem-loop structure (Lee et al., 2003).

In canonical pathway following nuclear processing, pri-miRNAs are exported to the cytoplasm which is mediated by exportin 5 (EXP5), a member of the nuclear transport receptor family (Krol et al., 2010). EXP5 recognizes the >14-bp dsRNA stem along with a
short 3’ overhang (1–8 nt) to bind. As with other nuclear transport receptors, EXP5 binds cooperatively to its cargo and the GTP bound form of the co-factor RAN in the nucleus, and discharges the cargo following the hydrolysis of GTP once it swims in the cytoplasm. Following export from the nucleus, pre-miRNAs are cleaved near the terminal loop by Dicer which contains a putative helicase domain, a DUF283 domain, a PAZ (Piwi–Argonaute–Zwille) domain, two tandem RNase-III domains and a dsRNA-binding domain (Ketting et al., 2001). Dicer positions the site of the second RNase-III cleavage on the stem of the miRNA precursors thereby generates mature miRNAs that range from 21 to 25 nucleotides (Carmell & Hannon 2004).

Figure 3.18. Biogenesis and diverse functions of miRNAs.
Adopted from Schwarzenbach et al., (2014).

Succeeding Dicer cleavage, the resulting RNA duplex is loaded onto an Ago protein to generate the effector complex, RNA-induced silencing complex (RISC) (Pellino et al., 2003). One strand of the ~22 nt RNA duplex remains in Ago as a mature miRNA called the guide strand, whereas the other strand (the passenger strand) is degraded. An RNA helicase activity is thought to mediate the unwinding and removal of the unselected strand of the
miRNA duplex. Dicer, TRBP (and/or PACT) and Ago proteins contribute to RISC assembly by forming a RISC loading complex (RLC). Now on this RISC complex identifies their corresponding target mRNAs to cleave or translational repression.

**MiRNAs - Mechanism of action**

![Diagram of miRNA mediated regulation](https://via.placeholder.com/150)

**Figure 3.19. Mechanism of miRNA mediated regulation.** Model showing the sequence complementarity between miRNAs (red) and the 3′-untranslated region (UTR) of target mRNA (blue) suppresses the protein synthesis machinery. RISC – RNA induced silencing complex. Adopted from Lin et al., (2005)

The miRNA recruited to the RNA-induced silencing complex (RISC) and regulates the production of protein-coding genes through diverse mechanisms. The interaction of miRNAs with the 3′ untranslated region (3′ UTR) of protein-coding genes is considered as the main mechanism, which leads to a decrease in target protein levels either by mRNA degradation or by translational repression. The perfect match of miRNA seed sequence and target mRNA’s UTR results in the degradation of the mRNAs whereas, the imperfect match repress the translation process thereby silencing the gene function. In plants it is shown to degrade the mRNAs whereas in animal it does not. Later studies have also suggested that miRNAs can interact with the 5′ UTR via complementarity and cause translational repression or activation of the targeted proteins (Lytle et al., 2007). Likewise, they can also target the coding sequence and suppress the translation of targeted genes (Tay et al., 2008). Moreover, some miRNAs can interact with regulatory protein complexes, and indirectly upregulate the translation of a target gene (Vasudevan et al., 2007).

MicroRNAs (miRNA) have recently been expanded to include in the known classes of genes of regulatory molecules that function as tumor suppressors and oncogenes family. miRNAs negatively regulate the solidity and translation of target messenger RNAs (mRNA) and have been involved in various cellular processes such as differentiation, cell-cycle control and apoptosis. It is obvious that loss- or gain-of-function of specific miRNAs subsidizes to cellular transformation and tumorigenesis, since the miRNA expression patterns are highly specific for cell-type and cellular differentiation status (Garzon et al., 2009). Incorporation of miRNA regulation in molecular pathogenesis of cancer will be essential to achieve a complete understanding of diseases. Similar to mRNAs, the miRNAs can also have oncogenic (oncomiR) or tumor suppressive (tumor suppressor miRNA) functions. Though there are huge number of miRNAs identified only few of them have been validated where the signature of some miRNAs exhibit disparity among different type of cancers.

MiR-21

MiR-21 is an oncogenic microRNA found to be overexpressed in many cancers including gastric cancer suggesting that it could be included as one of the diagnostic indicators of gastric carcinoma. A potential oncogenic role for miR-21 was discovered in a screen for abnormally expressed miRNAs in glioblastoma later which was found that to be expressed in higher levels in almost all tumors (Volinia et al., 2006). Aberrant expression of miR-21 may promote tumorigenesis by inhibiting apoptosis but no targets of miR-21 have yet been identified clearly, leaving the mechanisms of anti-apoptotic activity of this miRNA unexplained (Kent & Mendell, 2006). In general, expression patterns of miR-21 various carcinomas can promote cell growth suggesting that miR-21 might play a key role during the earliest-stages stomach carcinogenesis. Hence, miR-21 is considered one of the most important cancer specific miRNA that could serve as a good diagnostic marker for patients with gastric cancer (Chan et al., 2008).

MiR-145

Invasion and metastasis are key feature of malignant tumors that are responsible for 90% of cancer-related deaths MicroRNAs have been shown to have important role in suppressing tumor metastasis. MiR-145 is an interesting candidate that is deregulated in various cancers. In gastric cancer, downregulation of miR-145 level have been witnessed
in secondary metastasis of primary gastric cancers. Ectopic expression and loss-of-function of miR-145 have been observed to suppress gastric cancer cell migration and invasion. Moreover, N-cadherin (CDH2) has been proved to be a direct target of miR-145 where inhibiting N-cadherin could suppress tumor metastasis (Gao et al., 2013). The promoter of miR-145 have p53 binding elements which suggest that miR-145 can be regulated by p53 interaction. Similarly, the positive feedback between miR-145 and p53 have been identified which can impair the p53-murine double minute 2 (MDM2) feedback loop (Cui et al., 2014).

**MiR-148a**

MiR-148a is an interesting candidate involved in regulation of genes associated with DNA methylation which in turn control several different target genes and pathways involving tumor proliferation, invasion and metastasis. In gastric cancer, the miR-148a levels have been observed to be down-regulated in gastric cancer tissues (Xia et al., 2014). Thus it may serve as a novel biomarker for the diagnosis and as a new therapeutic target in gastric cancer. Zheng et al., (2011) proved that overexpression of miR-148a in gastric cancer cells can exert inhibitory effects of cell invasion and metastasis. The study also reported that miR-148a can regulate cancer invasion and metastasis by targeting ROCK, an essential effector kinase of Rho GTPases which plays a vital role in cellular migration. Downregulation of miR-148a in gastric cancer may enhance ROCK1 expression and likely to play a crucial role in development of a gastric neoplasia.

### 3.15. Competing Endogenous RNA

Competing endogenous RNAs are basically long non-coding RNAs that are able to sponge miRNAs and allow the mRNA evading from miRNA mediated regulation (Salmena et al., 2011). LncRNAs are the basic elements of the genome which were once thought to be a dark matter and play no biological role. Recently, the heap of research evidences shed light on these crucial players and now it has been established that they are potential regulators of genomic architecture. Similar to mRNAs, the LncRNAs can also have miRNA response elements (MREs) by which they absorb miRNAs that originally target the mRNAs. This endogenous competition between LncRNA and mRNA may be resulting in altered expression genes involved in various cellular processes and cancer (Ergun & Oztuzcu 2015). The previous mode, ‘miRNAs–RNAs’ regulation, now has been evolved to ‘RNAs–miRNAs–RNAs’ interaction.
Figure 3.20. Mechanism of competing endogenous RNA. (A) LncRNAs downregulation allows miRNAs binding to their target mRNA. (B) Overexpressed LncRNA competes with mRNA and sponge the targeting miRNA and allowing the mRNA free from miRNA mediated regulation. Adopted from Xia et al., (2014).

3.16. CeRNAs associated in gastric tumorigenesis

In recent decades, the mechanism of ncRNAs has attracted an adequate number cancer researchers. Increasing experimental evidences has opened up new windows to better understanding of ncRNAs and post-transcriptional gene regulation. In gastric cancer, the crosstalk between the lncRNAs and miRNAs are mainly focused where lncRNAs acting as CeRNAs are reported frequently. Recently, Li et al., (2016) reported that a large groups of lncRNAs have be associated with variety of cellular functions in cancer. Among them, HOTAIR, BC032469 and GAPLINC lncRNAs have been shown to act as ceRNA by sponging miR-331, miR-152, MiR-1207 and miR-211. These competitive binding could relieve HER2, hTERT and CD44 from their miRNA mediated regulation which ultimately results in apoptosis, migration and tumorigenicity in gastric cancer.

Amplification of H19 is such an example. H19 is an IncRNA that host a miR-675 coding sequence. The activation of H19 could also bring miR-675 which in turn triggers
molecular events. The H19 has been shown to be an oncogene activated by c-Myc that can regulate the p53 expression levels. Similarly the miR-675 could target mRNA of RB gene transcript. Hence, it makes clear that H19 overexpression could be favourable for cancer development.

Figure 3.21 LncRNAs regulate cell proliferation, cell cycle, apoptosis, invasion, migration, metastasis, and tumorigenicity. a. TINCR duplexes with KLF2 mRNA. b. MALAT1 influences splicing. c. GHET1 and c-Myc stability. d. H19 generates miR-67 and .H19 binds to miR-141 e. TUSC7 negatively targets miR-23b. f. miR-148a reduces hypermethylation of MEG3 g. MIR-49A/MIR-449A epigenetically silence ANRIL. Adopted from Li et al., (2016).

In gastric cancer, overexpressed H19 have been shown to promote the proliferation of gastric cancer cells where its co-transcript miR-675 have been shown to target RUNX1 (Li et al., 2014). This suggest that interaction between H19/miR-675/RUNX1 could be considered as novel candidates that organise gastric tumorigenesis. Yan et al., (2014) reported that MEG3 can sponge the miR-148a and allows its downstream target DNMT1 to overexpress in gastric cancer. The study also highlighted that the MEG3 levels are elevated in gastric cancer cell line sand tumor tissues. Similarly, lncRNA AC130710 was
shown to be regulating miR-129-5p where HOTAIR is another lncRNA demonstrated to suppress miR-331-3p.

3.17. CeRNA network involved in gastric cancer

Competitive endogenous RNAs cross-regulate each other through sequestration of shared microRNAs and form some complex regulatory networks based on the dominant expression pattern at certain environmental stimuli. In general, numerous long non-coding RNAs (lncRNAs) and pseudogenes have been predicted and experimentally confirmed to crosstalk with each other through inherent miRNA binding sites alternatively termed as miRNA response elements (Poliseno et al., 2010 & Seitz, 2009). There are myriad number of transcripts which harbor one or many MREs that enable the binding of targeting miRNAs (Salmena et al., 2011). This feature ultimately makes each and every ceRNA to be a part of a complex network proposed as “ceRNETs”. One such exemplary example has been verified with phosphatase and tensin homolog (PTEN) and its ceRNA VAMP (vesicle-associated membrane protein)-associated protein A (VAPA). Though many candidates are identified every day, the degree of ceRNA networks remains unclear which might be virtuously influenced by interim molecular signatures initiated by genomic events.

To overcome this issue, many qualitative prediction and quantitative validations are on pipeline to create ceRNET models. These are all achieved by a bioinformatic algorithms with mathematical integration of molecular profiles collected from experimental datasets. These models are promising and further investigations are needed to validate their critical interactions. Furthermore, cellular and molecular events can be characterized by blend of several ceRNAs that could form subnetworks which are basic structures of larger networks involved in multiple processes of an animal system. These subnetworks are often centralised with novel ceRNA that are connected with one or more miRNAs which ensure the minimal fluctuations of a tightly regulated network. This indicates that loss of one ceRNA might alter the skeleton of a subnetwork that further leads to intense effects on basic components and alteration in the architecture of larger network (Ala et al., 2013).

The competing endogenous RNA network has three distinct features namely: (1) relative concentration of targeting miRNAs (2) tissue specificity and subcellular localization of ceRNA (3) inequality of miRNA response elements (MREs) on ceRNAs. All these three characteristics can potentially influence the regulatory mechanism of
CeRNA in different ways (Hu et al., 2015) In general, the biological activity of ceRNA is dependent on the balancing between miRNAs its target lncRNA or pseudogenes. The imbalance of these regulatory RNAs could elude a variety of cellular malfunctions that eventually promote diseases like cancer. In addition to that, 3′ UTRs lengthening or shortening initiated by alternative polyadenylation signals, seed sequence polymorphisms or mutations have also been reported as rheostats for ceRNA interactions (Li et al., 2014). Global miRNA expression profiling studies have highlighted that both tumor suppressive or oncogenic miRNAs are unconditionally exhibit heterogeneity among different cancers including gastric cancer. These complex networks can be connected with multiple cellular and molecular events in gastric tumorigenesis and complexity of stomach cancer.

Most recent bioinformatic and experimental studies on gastric cancer have established that differential expression pattern of lncRNA and deregulated miRNA profile could be constructed into a ceRNA network. For example, FER1L4 is an lncRNA which has been found to be involved in sponging miR-106a-5p. This lncRNA competitively binds with miR-106a-5p by sharing common MREs between the RB1 transcript (Xia et al., 2014). The regulatory interaction between FER1L4 and miR-106-5p could facilitate escape of RB1 from miRNA mediated regulation and its upregulation. This has been verified with luciferase functional studies which strongly suggest that miR-106a-5p can serve a core element where lncRNAs like FER1L4, GACAT1 and H19 could act as ceRNAs to regulate the expression level of many cancer associated genes like PTEN, RB1, RUNX1, VEGFA, CDKN1A and E2F1. Another recent study has found that HER2 (human epidermal growth factor receptor 2) is positively regulated by endogenous competition of HOTAIR through the suppression of miR-331-3p (Liu et al., 2014 & Deng et al., 2015). This interaction was evidenced in restoration of HER2 in gastric cancer where the ectopic expression of HOTAIR was reported to enhance invasion and metastatic potential of cancer cells. Similarly, the elevated levels of miR-331-3p was observed to be correlated with depletion of HOTAIR.

Even though the ceRNA concept is under developing stage, the present knowledge is growing enormously which attracted much attention. Advanced genome analysis platforms and high-throughput analytical methods has restructured the data quality that has maximized the enhanced ceRNET prediction tools and validation methods. Similarly, large
scale data collection and meta-analysis have revealed new novel candidates with literature evidences. These considerable volume of data are deposited in public databases (Table 3.4).

<table>
<thead>
<tr>
<th>Database</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>ceRDB</td>
<td><a href="http://www.oncomir.umn.edu/cefinder/">http://www.oncomir.umn.edu/cefinder/</a></td>
</tr>
<tr>
<td>lnCoDB</td>
<td><a href="http://gyamot-beta.com/Lncodb/">http://gyamot-beta.com/Lncodb/</a></td>
</tr>
<tr>
<td>Lin2GO</td>
<td><a href="http://www.bioinfo.tsinghua.edu.cn/~liuken/Lin2GO/index.html">http://www.bioinfo.tsinghua.edu.cn/~liuken/Lin2GO/index.html</a></td>
</tr>
<tr>
<td>starBase v2.0</td>
<td><a href="http://starbase.sysu.edu.cn/">http://starbase.sysu.edu.cn/</a></td>
</tr>
<tr>
<td>miRcode</td>
<td><a href="http://www.mircode.org/">http://www.mircode.org/</a></td>
</tr>
<tr>
<td>ChIPBase</td>
<td><a href="http://deepbase.sysu.edu.cn/chipbase/">http://deepbase.sysu.edu.cn/chipbase/</a></td>
</tr>
</tbody>
</table>

**Figure 3.4. CeRNA databases available online**

Taking all these into consideration, ceRNA interplay has been proven a pervasive mechanism in many carcinogenesis including gastric cancer which also clearly establishes that ceRNA interactive network are crucial for gastric tumorigenesis and development. Hence, it is anticipated that identification of new ceRNA networks could be useful for the identification of key molecules that can be used for both diagnosis and prognosis in gastric cancer.