1. INTRODUCTION

Gastric cancer is a serious global health issue that stands as the third most frequent cause of cancer death and fifth most common cancer after lung, breast, colorectum and prostate cancers. As per the GLOBOCAN 2012 update, gastric cancer constitutes 6.8% of the total cancer incidences worldwide. The occurrence of the disease varies geographically where nearly half of the global incidence is observed in Asia especially in countries like China and other East Asian nations including Japan, South Korea and Taiwan (Ferlay et al., 2015). Other Asian countries including India, Pakistan, and Thailand show moderately low incidence. In contrast, the developed countries from western part of the world show slightly decreased incidence for the past few decades. The international observations suggest that the higher incidence could be caused by either genetic or environmental or both. Similarly, the lower incidence could have been resulted by recent changes in the environmental or other co-operating factors associated with gastric carcinogenesis. Though the incidence is quite inconsistent, majority of the cases (>90%) are often diagnosed at advanced stage which delimit the clinical treatment options and therapeutic strategies. Recent advances in medical diagnosis have revolutionized the early screening and management of the disease where only 70% cases are cured with surgical therapy. While surgical therapy is the most promising treatment modality, the disease free survival rate is moderately less than other cancers. In addition to that, lack of non-invasive diagnostic method keeps the early identification of this disease defiant for decades. Hence, the eradication of gastric cancer remains a great challenge in both clinical and diagnostic arena.

Gastric tumorigenesis, a multistep process, starts from a simple atrophic gastritis that subsequently develops into a metaplasia which finally transformed into a cancerous neoplasm. Pathological evidences highlight that Gastric Adenocarcinoma which arises at glandular epithelium of stomach is the major type of gastric cancer which accounts for approximately 95% of the total cases identified (Carcas, 2014). The combination of both the environmental and genetic factors is considered as equally important for the development of gastric tumorigenesis where the causative role differs for two subtypes namely (i) Diffuse or cardia type (ii) Intestinal or non-cardia type (Lauren et al., 1965). The overall incidence is mainly occupied by the diffuse type which remains stable with slight increasing rate of incidence among different populations. Gastric Adenocarcinoma exists as an important clinical morbidity since it is mostly developed
by normal human aging process. Also, it is an extremely heterogeneous disease which displays wide spectrum of etiological traits that includes *Helicobacter Pylori* infection, tobacco, alcohol, cured meat and salted food. Apart from lifestyle associated risk factors, *Helicobacter Pylori* infection has been classically known as an important risk factor which has been classified under type I carcinogen (Amieva *et al.*, 2016). Similarly, the food habits are strongly associated with stomach cancer. Unlike the environmental risk factors, the genetic make-up of an individual is thought to be an independent risk factor for the gastric malignancies in different ethnic groups all over the world.

The development of gastric carcinogenesis is achieved by various stages where each phase is enabled with detrimental genomic modifications. Identification of these aberrant genetic and epigenetic alterations could be effectively applied to improve the earlier diagnosis of gastric cancer. In addition to that, overall survival of patients can be maximized with tailored therapeutic strategies that are based on molecular genetics (McLean *et al.*, 2014). To date, extensive cancer research studies have been carried out to address the unique cellular and molecular genetic events involved in the gastric cancer development and progression. In connection to this, altered gene expression level of various genes has been frequently reported to be associated with glandular epithelial malignancies of stomach. However, identification of a biomarker or a molecular classifier remains unexplored for this multifactorial disease.

Advances in next generation sequencing and whole transcriptome analysis have revealed wide array of genes that are frequently involved in gastric cancer. Many of these genes are expressed throughout abnormal tumorigenic cascade where the expression level of their mRNAs are often linked with different clinical stages and pathological grades. Similarly, these abundantly expressed genes could also define a unique molecular signature in gastric adenocarcinoma (Wang *et al.*, 2011). Hence, these genes play a central role in signaling pathways essential for cell growth and proliferation. Identification of gastric cancer specific genes require consistent scientific reports with adequate experimental support. Existing scientific evidences highlight that a set of candidate genes namely EGFR, FGFR2, KLF4, DNMT1 and AGO4 are involved in the gastric tumorigenesis. These genes are basically the cell surface receptors, transcription factor and proteins involved in DNA and RNA binding. The protein product of these genes are core partner of major cellular events like cell-cell
communication, transcriptional activation, epigenetic modification and RNA processing.

Non-coding RNAs are the recent attractions in genome biology which establish a largest class of transcripts that occupies ~80-90% of a human genome. These RNA transcripts lack the potential to be translated into a protein which are abundant in almost all tissues. Based upon their size and biological functions, these ncRNAs have been further classified into many types including microRNA (miRNA), long non-coding RNA (LncRNA), and Transcripts of unknown function (TUFs). Unlike the messenger RNAs (mRNA), the ncRNAs could interact with DNA, RNA and Protein. In recent decades, experimental evidences demonstrate that noncoding RNA play multifunctional role in almost all major genomic events like transcriptional regulation, maintenance of protein level and genome integrity (Song et al., 2013). Recent findings make these ncRNAs inevitable players in cancer evolution. It is well established that ncRNA interaction with mRNA transcribed from tumor suppressors or oncogenes could modify the protein level in various cancers where mutational alterations are rare.

MicroRNAs (miRNAs) are a small endogenous non-coding RNAs made up of ~22 nt in length. They are highly conserved and present in viruses, plants and animals (Bartel, 2004). To date, nearly 3,786 miRNA candidates have been validated where all of them possess a unique seed sequence made up of 8 nucleotide which could be complementary to the 3’-untranslated region of target mRNAs. These miRNA interactions could be potentially influencing the gene expression of the target mRNA at post transcriptional level. The miRNA mediated gene regulation is considered as more critical event since they directly intervene with the mRNA translation and modify the protein level. These tiny RNAs have a very big role in almost all cells types of our body. MiRNA dysregulation are frequently reported in all type of cancers where the miRNA are massively transcribed cancerous cells.

In gastric adenocarcinoma, many reports strongly suggest that expression level of miRNA could be considered as biomarkers for the molecular classification of tumor subtype, aggression and progression. Among the large number of miRNAs, a very few miRNA candidates has been already reported to be overexpressed in gastric malignancies. MiRNA candidates like hsa-miR-21, hsa-miR-145, and hsa-miR-148a have been reported to be deregulated in gastric cancer subjects from different
populations. The miR-21 is a well-established oncogenic miRNA which has been observed to be implicated in downregulation of various tumor associated genes. The other two candidates miR-145 and miR-148a are frequently reported to be involved in many cancers where the miR-145 plays a major role in stemness and miR-148a is linked with DNA methylation.

Long non-coding RNAs are >200 bp sized transcripts which are usually not translated into proteins. The LncRNAs are key players in genome biology which help to maintain the genome architecture of the normal cells by various processes including epigenetic modifications. They are abundantly expressed and play a very important role in different biological processes like development, aging and cancer. Previously, the loss of genome integrity in cancer cells was poorly understood and now it has been well established that the alteration in the expression level of these lncRNAs could modify cellular characteristics and can even result in tumorigenesis at certain conditions. Increasing number of scientific reports strongly suggest that these lncRNAs play diverse role in gastric cancer (Fang et al., 2015). The current experimental evidences point out that lncRNAs like H19, GAS5, PTENP1-AS, and MEG3 show differential expression among stomach cancer patients. These observations could be used to define the molecular background of gastric carcinogenesis that may eventually lead to identify the central player that can be possibly used as diagnostic biomarker or therapeutic target in patients.

Apart from serving as a backbone for maintaining genome integrity, the lncRNAs can also modify the genomic events and cellular homeostasis through dynamic interaction with other type of RNA transcripts like miRNA and mRNA. A recently postulated theory sheds light on competitive nature of lncRNA with mRNA by modulating the miRNA mediated regulation. The lncRNAs have been identified to have one or more miRNA binding sites also known as microRNA responsible elements (MRE) similar to the target mRNAs. This enables the lncRNAs to sponge the miRNAs and allowing the target mRNAs to escape from miRNA regulation resulting in remarkable upregulation of genes. The lncRNAs that potentially absorb the miRNAs and capable of blocking the miRNA mediated activity are generally termed as “competitive endogenous RNA” or CeRNA. The miRNA binding competitiveness between the protein coding mRNA and non-coding lncRNA can be alternatively regulated by each other’s expression. This could eventually modify the mRNA turnover
resulting in increased or decreased protein level (Salmena et al., 2011). Moreover, the extensive interaction among different type of CeRNA could form a larger regulatory networks that are known as ceRNA networks (ceRNETs). These regulatory RNA networks are more promising shed lights on the origin of cancer and development. To support the ceRNA hypothesis, a few recent reports have clarified that IncRNAs could really serve as CeRNA in multiple cancers including gastric cancer (Xia et al., 2014). Identification of ceRNA could be promising to understand the clinical importance of post transcriptional regulatory interactions between protein coding mRNAs and non-coding RNAs that eventually orchestrate the molecular mechanisms of gastric tumorigenesis and clinical manifestations. Hence, this thesis was intended to identify few lncRNA that could act as ceRNA in gastric cancer.

The experimental study was conceived with the aim of validating the ceRNA hypothesis by identifying IncRNAs which could act as a ceRNA in gastric cancer. To achieve this, the study was designed with four different steps namely (1) Expression profiling of cancer associated LncRNAs (2) profiling of gastric cancer specific genes associated with diagnosis and prognosis (3) Expression analysis of differentially expressed miRNAs (4) Identification of lncRNA that form a ceRNA network in gastric cancer. Each step involved in the study is so specific for different type of RNA where their expression levels were individually assessed by various mode of high-through put quantitative PCR. To profile the cancer associated lncRNAs, we selected a total number of 18 cancer associated IncRNAs from the previous reports. Their expression levels were quantified in tumors and normal tissue samples by a customized Taqman probe based lncRNA quantification method. Similarly, 5 gastric cancer specific genes and 3 different miRNAs were selected from existing scientific literature and quantified by relative quantitation method using qRT-PCR. The quantitative results obtained for each class of RNA were statistically analyzed and the integrated network analysis of IncRNA, miRNA and mRNA revealed the existence of ceRNA network that might be critically involved in tumorigenesis of gastric cancer patients of Indian ancestry.