Am J Stem Cells 2014;3(2):63-73
www.AJSC.us /ISSN:2160-4150/AJSC0001489

Review Article

Emerging paradigms of long non-coding RNAs in gastrointestinal cancer

Hui-Juan Jiang, Shuang Wang, Yanqing Ding

Department of Pathology, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, China

Received July 21, 2014; Accepted August 7; 2014; Epub September 5, 2014; Published September 15, 2014

Abstract: A large number of long non-coding RNAs (lncRNAs) have been discovered by genome-wide transcriptional analyses. Emerging evidence has indicated that lncRNAs regulate gene expression at epigenetic, transcription, and post-transcription levels, are widely involved in various pathobiology of human diseases, and may play an important role in the biology of cancer stem cells. Alterations of specific lncRNAs have been revealed to interact with the major pathways of cell proliferation, apoptosis, differentiation, invasion and metastasis in many human malignancies, such as gastrointestinal cancer. This review summarizes the current understandings in biological functions and implications of lncRNAs in gastrointestinal cancer.

Keywords: Long non-coding RNAs (lncRNAs), gastrointestinal cancer, tumor progression

Introduction

For many years, the so-called central dogma of molecular biology has been dominant in biomedical fields, including cancer research [1], which assumes that the information flow in a cell is uniquely directed from DNA to RNA to protein, and that proteins are ultimately responsible for cell phenotype. However, recent advances in genomics technology, such as tiling arrays and RNA deep sequencing, have led to the explosive discovery of pervasive transcription activity in most of the eukaryotic genomic regions that were once considered "junk DNA" or "deserts" [2-4]. The studies have revealed that approximately 97% of eukaryotic genomes are transcribed, whereas only 3% of the genome encodes proteins [5]. This suggests that a large proportion of the genome produces an unexpected plethora of RNA molecules that have no protein coding potential. These are collectively called noncoding RNAs (ncRNAs). Based on their size, ncRNAs include long (or large) non-coding RNAs (lncRNAs) and short ncRNAs, including ribosomal-RNAs (rRNAs), ransfer-RNAs, small interfering RNAs (siRNAs), Piwi-RNAs (piRNAs), as well as microRNAs. The human and other mammalian genomes produce thousands of lncRNAs [6-8]. To date, only a handful of lncRNAs have been studied functionally and/or mechanistically. However, lncRNAs have recently drawn intense attention with the bright perspective that they may represent a new regulatory layer in the complexity of mammalian gene regulatory networks underneath a wide range of pathobiology of human diseases. Evidence shows that lncRNAs could play crucial roles in development and progression in gastrointestinal cancer. Here, we provide an overview on lncRNAs and discuss the current and future research that may shed further light on understanding the functional role of lncRNA in gastrointestinal cancer.

Basic biology of lncRNAs

The definition of lncRNAs

LncRNA is conventionally defined as an endogenous and a non-protein-coding RNA molecule longer than 200 nucleotides, which lacks protein-coding capability [9]. This length is a convenient cut-off to exclude small RNAs in a RNA purification procedure, but include pseudogenes, microRNA precursors as well as RNAs

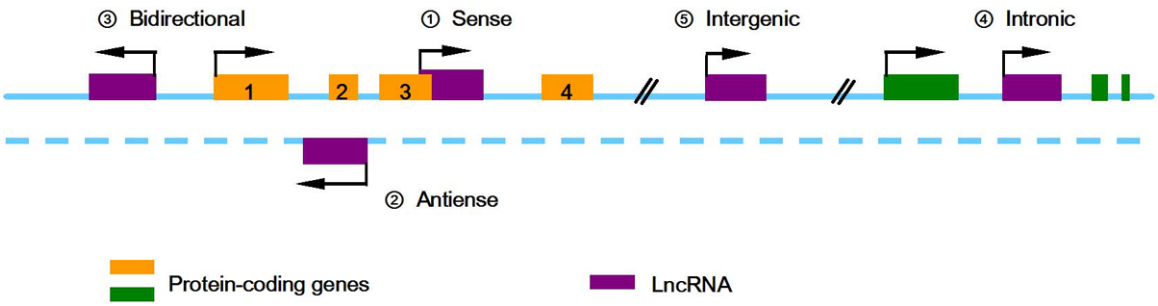

Figure 1. LncRNAs can be divided into five categories according to their location relative to nearby protein-coding genes. ① Sense lncRNAs sequence overlaps with the sense strand of a protein-coding gene; ② Antisense lncRNAs are transcribed in the opposite direction of protein-coding genes, and overlap at least one coding exon; ③ Bidirectional lncRNAs are transcripts that initiate in a divergent fashion from the promoter of a protein-coding gene; ④ Intronic lncRNAs are lncRNAs that initiate inside of an intron of a protein-coding gene in either direction and terminate without overlapping exon.

that interact with epigenetic effectors and splicing factors. LncRNAs are strikingly similar to mRNA: they are RNA polymerase II transcripts, which are capped, spliced and polyadenylated, yet do not function as templates for protein synthesis.

Structures of lncRNAs

Although lncRNAs constitute a large fraction of the transcriptome, only a few lncRNAs have been structurally annotated and found to share some primary sequence features. The sequenced features in lncRNAs have been obtained by analyzing 204 functional lncRNAs and their splicing variants, including paucity of introns, low GC content, poor start codon and ORF contexts [10]. In addition, the presence of motifs embedded in the lncRNA primary sequence enables lncRNAs to be specifically associated with DNA, RNA, and protein. Some lncRNAs alternate miRNA binding sites to regulation expression levels of protein-coding genes [11]. The large- or small-scale mutations in the lncRNA primary sequence are highly correlated with diseases [12, 13].

LncRNAs lack conservation in many cases even among closely related species. Many lncRNAs have a significant secondary structure which is critical for specific binding and function [14]. For example, lncRNA SRA has a complex structural architecture, which is organized into four distinct domains, with a variety of secondary structure elements to interact with a variety of proteins [15]. Additionally, non-canonical end maturation of MALAT1 lncRNA involves a cloverleaf secondary element at its 3'-end [16].

Many lncRNAs have multiple alternatively spliced forms. For example, lncRNA PVT-1 produces a wide variety of spliced noncoding RNAs as well as a cluster of six annotated microRNAs: miR-1204, miR-1205, miR-1206, miR-1207-5p, miR-1207-3p and miR-1208 [17]. The NCBI AceView database has revealed at least 10 alternative RNA transcripts derived from the colorectal neoplasia differentially expressed (CRNDE) locus [18]. Alternative splicing is now recognized to contribute to the pathogenesis of many diseases, including cancers [19]. Thus, structural architecture and identification of splice variants for lncRNAs may contribute to a better understanding of the mechanisms responsible for lncRNAs function.

The origins of LncRNA

There are five possible sources from which the lncRNA is resulted: (1) a protein-coding gene acquires frame disruptions and is transformed into a functional noncoding RNA that incorporates some previous coding sequence. The Xist lncRNA originated by undergoing a metamorphosis from a previous protein-coding gene while incorporating transposable element sequence; (2) following a chromosome's rearrangement, two untranscribed and previously well-separated sequence regions are juxtaposed and give rise to a multi-exon noncoding RNA. A dog noncoding RNA (supported by ESTsBM537447, C0597044, and DN744681)

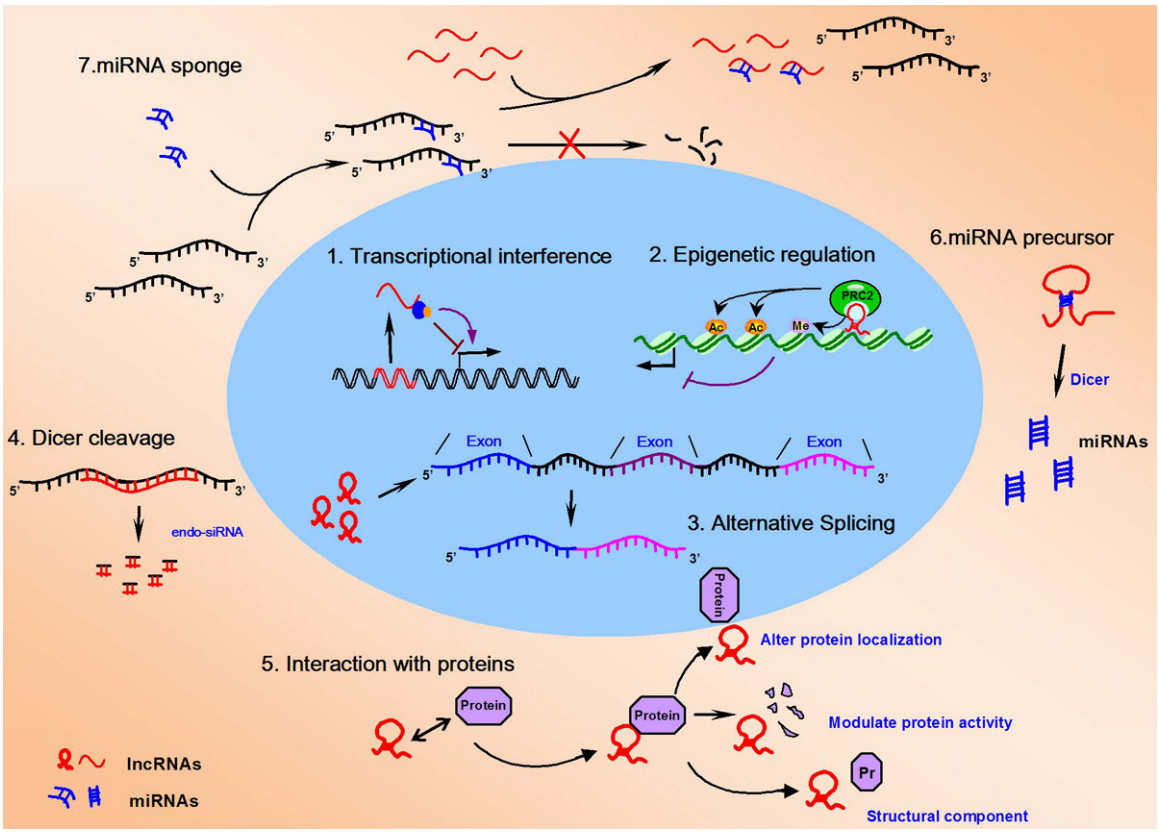

Figure 2. Schematic illustration for cellular functions of lncRNAs (red). 1. lncRNAs can either negatively or positively affect expression of the coding gene (black) by transcriptional interference; 2. lncRNAs can recruit chromatin modification complex to regulate coding gene expression by inducing chromatin remodeling and DNA methylation; 3. lncRNAs can change alternative splicing of various pre-mRNAs; 4. antisense lncRNA can pair to their specific sense mRNA, generating endo-siRNAs; 5. lncRNAs can interact with proteins to influence protein activity, to alter protein localization, or to modulate structural component; 6. lncRNAs can be processed to yield miRNAs; 7. lncRNAs can also act as "miRNA sponge" to influence the mRNA expression mediated by miRNA.

appears to be yielded following such a lineage-specific change; (3) duplication of a noncoding gene by retrotransposition generates either a functional noncoding retrogene or a nonfunctional noncoding retropseudogene; (4) neighboring repeats within a noncoding RNA have their origins in two tandem duplication events; and (5) insertion of a transposable element gives rise to a functional noncoding RNA [20].

The classification of LncRNAs

On the basis of their position relative to protein-coding genes, lncRNAs can be divided into five broad categories (**Figure 1**): (1) sense; (2) antisense, when overlapping one or more exons of another transcript on the same, or opposite strand, respectively; (3) bidirectional, when the expression of a lncRNA and a neighboring coding transcript on the opposite strand is initiated in close genomic proximity; (4) intronic, when it is derived from an intron of a second transcript (although these, as noted above, sometimes may represent pre-mRNA sequences); and (5) intergenic, when lncRNA lies within the genomic interval between two genes [21].

It has been indicated that lncRNA classification may reflect functional characterization. Clark et al [22] determined the half-lives of approximately 800 lncRNAs in the mouse Neuro-2a cell line and revealed that intergenic and cis-antisense lncRNAs were more stable than those derived from introns. Additionally, they found that many human large intergenic noncoding RNAs could be associated with chromatin-modifying complexes and affect gene expression [23]. Therefore, the classification of lncRNAs can provide crucial pieces of information for studying the potential function of lncRNAs in some extent.

Table 1. lncRNAs that have been or might be linked to gastrointestinal cancers

lncRNAs	Expression	Cancers	Emerging roles	References
HOTAIR	Up-regulation	CRC	Metastasis, prognosis	31,58
		GC	Metastasis	32,59,60
CCAT1	Up-regulation	GC	Proliferation	34,52
H19	Up-regulation	GC	Diagnosis, apoptosis and proliferation	33,44,53,54
CCAT2	Up-regulation	CRC	Microstellite stability	36
p21	Down-regulation	CRC		37
BM741401	Down-regulation	GC	Prognosis,metastasis	38,
CCAT1	Up-regulation	CRC	Diagnosis	39
AC096655.1-002	Down-regulation	GC	Diagnosis	40
MEG3	Down-regulation	GC	Prognosis, apoptosis	41
CRNDE	Up-regulation	CRC	Diagnosis	19
PVT1	Up-regulation	CRC	Proliferation, apoptosis	49
MALAT1	Up-regulation	CRC	Metastasis	56
PCAT-1	Up-regulation	CRC	Metastasis	62

The functions of lncRNAs

Although a functional lncRNA known as Xist was discovered and characterized in the early 1990s [24-26], only a small percentage of lncRNAs have been studied in detail. Increasing numbers of lncRNAs have been shown to function in development and participate in a wide variety of molecular genetic and cellular processes, such as chromosomal dosage compensation, control of imprinting, chromatin modification, chromatin structure, transcription, splicing, translation, cellular differentiation, integrity of cellular structures, cell cycle regulation, intracellular trafficking, reprogramming of stem cells and the heat shock response [27]. **Figure 2** shows how lncRNA functions at molecular level: (1) transcription from an upstream noncoding promoter affect expression of the downstream gene; (2) It affects the expression of the downstream gene by inhibiting RNA polymerase II recruitment or inducing chromatin remodeling; (3) an antisense transcript is able to hybridize to the overlapping sense transcript and block recognition of the splice sites by the spliceosome, thereby resulting in an alternatively spliced transcript; or (4) alternatively, hybridization of the sense and antisense transcripts can allow Dicer to generate endogenous siRNAs. By binding to specific protein partners, a noncoding transcript can modulate the activity of the protein; (5) serve as a structural component that allows a larger RNA-protein complex to be formed or altered where the protein localizes, or to modulate proteins activity in the cell; (6) lncRNAs can be processed to yield small RNAs, such as miRNAs, piRNAs, and other less well-characterized classes of small transcripts [28]; and (7) lncRNAs can also act as "miRNA sponge" to influence the mRNA expression mediated by miRNA.

The relation between gastrointestinal cancer and lncRNAs

Although only a small percentage of human lncRNAs have been characterized so far. Notably, it has been shown that the roles for lncRNAs as drivers of tumor suppressive or oncogenic functions have appeared in diverse cancer types [29]. Currently, lncRNAs have been found to be deregulated in gastrointestinal cancer, and elucidated to influence the hallmarks of cancer, including cell proliferation, apoptosis, invasion and metastasis (**Table 1**).

Aberrant expression of lncRNAs in gastrointestinal cancer

During the past decade, it has been revealed several examples of differentially expressed

lncRNAs in cancer, some of which contributes to neoplasia [30]. Experiments with lncRNA microarrays to search for abundantly expressed lncRNAs in gastric cancer have revealed 135 lncRNAs with differential expression levels between tumor and non-tumor tissues [31]. The well-studied lncRNA, HOTAIR, is identified to be highly expressed in colorectal cancer (CRC) and gastric cancer [32, 33]. Yang et al [34, 35] revealed that lncRNA H19 and colon cancer associated transcript 1 (CCAT1) levels were markedly increased in gastric carcinoma tissues as compared with normal tissues, suggesting a critical role in the progression of gastric carcinoma. LncRNA CCAT2 is highly overexpressed in microsatellite-stable CRC [36].

Nevertheless, some lncRNAs exhibit down-regulated expression level in gastrointestinal cancers. LincRNA-p21 expression level was increased by elevated wild-type p53 induced by nutlin-3 in HCT-116 CRC cells. lincRNA-p21 was significantly lower in CRC tumor tissue as compared with the paired normal tissues [37]. Park et al. [38] performed RNA-seq experiments to compare gastric cancer with normal tissues and found differentially expressed transcripts in intergenic regions. They identified 31 transcripts, including BM742401 which was downregulated in cancer [38].

The role of lncRNAs in the diagnosis and prognosis prediction of gastrointestinal cancer

LncRNAs are found in the nucleus, cytoplasm or both. Intriguingly, many lncRNAs have restricted tissue- and cancer- specific expression patterns, which suggest potential applications in diagnostic and/or prognostic evaluation [6, 7, 23]. As for gastrointestinal cancer, some lncRNAs may be useful as novel potential biomarkers for diagnosis and/or prognosis prediction. CCAT1 lncRNA was found to be expressed in CRC tumors, but not in normal tissues, suggesting that CCAT1 is a powerful diagnostic parameter for the specific identification of CRC [39]. AC096655.1-002, for example, can be served as a diagnostic marker and predictor for cancer progression for gastric cancer patients [40]. AC096655.1-002 was significantly downregulated in gastric cancer tissues, and the sensitivity and specificity of the AC096655.1-002 in the diagnostic of gastric cancer was 0.513 and 0.872. In advanced gastric cancer stages, the low expression of AC096655.1-002 was associated with distant metastasis, lymph node metastasis, depth of invasion and poor survival. Additionally, the downregulated expression of lncRNA BM742-401 was correlated with poor survival in gastric cancer patients [38]. Similarly, the patients with lower expression of maternally expressed gene 3 (MEG3) had a significantly poorer prognosis than those with higher MEG3 expression in gastric cancer [41].

Moreover, similar to circulating miRNAs, some lncRNAs are demonstrated to be present in body fluids such as blood and urine, and can be detected by PCR [42, 43]. CRNDE lncRNA splice variants were up-regulated in neoplastic colorectal tissues. The expression level of CRNDE-h transcript in the plasma of CRC patients was 5.5 times greater than that of the healthy individuals. The expression level of CRNDE-h alone revealed a sensitivity of 87% and specificity of 93% for predicting the presence of CRC [18], suggesting that CRNDE transcripts may have clinical utility in screening and diagnosing CRC. In gastric cancer, plasma H19 level was found significantly higher in patients than healthy controls and reduced in postoperative samples [44]. The detection of circulating lncRNAs may provide new complementary tumor markers for gastric cancer, although the exact mechanism of the release of lncRNAs into body fluids remains elusive. It was found that lncRNAs were relatively stable in plasma samples, and were protected from the severe conditions tested by some mechanisms [44]. Given this specificity and accessibility, lncRNAs may be the biomarkers superior to many current protein-coding biomarkers.

The role of lncRNAs in the proliferation of gastrointestinal cancer

One of the most prominent characteristics of a cancer cell is its ability to achieve unlimited growth in the absence of external stimuli [45]. The small ncRNA including well-documented microRNAs receive the most attention and are shown to play many important roles in cancer proliferation [46, 47]. In several types of cancer, including CRC and gastric cancer, the expression of lncRNAs has led to promote or repress cell proliferation. Amplification of 8q24 is one of the most frequent events in a wide variety of malignant diseases including CRC [48]. PVT-1, which encodes a lncRNA, is mapped to chromosome 8q24 [7]. Takahashi et al found that 8q24 copy-number gain pro-

moted PVT-1 expression, and that aberrant expression of PVT-1 was of significance and accompanied by genomic alteration in CRC [49]. Reducing PVT-1 expression by siRNA resulted in a significant loss of CRC cell proliferation. In CRC cells with knockdown of PVI-1, the transforming growth factor beta (TGF-β) signaling and apoptosis signals were significantly activated [49]. Another example of lncRNA involved in proliferation is lincRNA MEG3, a 1721 bp noncoding RNA. MEG3 gene is an imprinted gene belonging to the imprinted DLK1-MEG3 locus located at chromosome 14q32.3 in human. Evidence elucidated that MEG3 is expressed in normal tissues while its expression is lost in an expanding list of primary human tumors and tumor cell lines [50, 51]. In gastric cancer, Sun et al. performed Hoechst staining analysis for tumor cells with ectopic over-expression MEG3 and illustrated that enforced expression of MEG3 significantly induced apoptosis in vitro, whereas the inhibition of MEG3 expression promoted the proliferation of gastric cancer [41]. Besides, another study showed that CCAT1 was also upregulated in gastric carcinoma in comparison with adjacent normal gastric tissues, and its abnormal expression promotes the proliferation of gastric carcinoma cells [52]. Furthermore, one E-box element located at the promoter region of CCAT1 and c-Myc could regulate both the promoter activity and expression of CCAT1 through direct binding to the E-box element. Enforced expression of c-Myc increased the expression of CCAT1. On the contrary, the inhibition of c-Myc also decreased the expression of CCAT1 correspondingly [52]. The up-regulation of H19 was found to contribute to proliferation of gastric cancer cells [53]. It was also indicated that H19-derived miR-675 could modulate human gastric cancer cell proliferation by targeting tumor suppressor RUNX1 [54].

The relation between lncRNAs and metastasis in gastrointestinal cancer

Increasing evidence has demonstrated that cancer patients die from metastases instead of the primary tumor. Thus, it is still critical for us to find novel biomarkers to predict the possibility of metastasis. Interestingly, many lncRNAs are consistently associated with clinical parametric indicatives of metastasis in a wide spectrum of tumor types [55]. A notorious example of such an oncogenic lncRNA is metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), which was initially found over-expressed in lung cancer metastases [56]. In our laboratory, Xu et al. [13] identified that a fragments (6918 nt - 8441 nt) located at the 3' end of MALAT-1 played a pivotal role in the invasion and metastasis of CRC cells. Analogous to MALAT1, homeobox transcript antisense RNA (HOTAIR) represents another lncRNA that is strongly associated with metastatic progression. HOTAIR is a long non-coding RNA that was identified from a custom tilling array of the HOXC locus. HOTAIR can simultaneously interact with both polycomb repressive complexes 2 (PRC2) and LSD1/CO-rest, which catalyze histone H3K27 trimethylation and H3K4 demethylation, respectively. When over-expressed, HOTAIR targets these repressive complexes to inhibit HOXD gene expression, and promotes tumor invasion and metastasis in breast cancer [57]. In colorectal cancer (CRC), HOTAIR expression level was found higher in CRC tissues than that in the corresponding noncancerous tissues. High expression levels of HOTAIR correlated with the presence of liver metastasis, and CRC patients with a high HOTAIR expression level also had a worse prognosis than those with a low HOTAIR level [58]. In gastric cancer, it was found that the expression level of HOTAIR was significantly higher in gastric carcinoma lesions as compared to non-cancerous lesions, and HOTAIR might contribute to distant metastasis and/or peritoneal dissemination rather than direct invasion to neighboring organs [59]. Furthermore, Xu et al [60] performed the transwell matrigel invasion assays and showed that down-regulation of HOTAIR by siRNAs caused a significant decrease in the cell invasiveness. While representing only a few examples of an increasing body of literatures, MALAT1 and HOTAIR provide a solid rationale for developing more lncRNA-based bests aiming at assessing the pro-metastatic potential of gastrointestinal cancer.

In addition to above-discussed lncRNAs that promote metastasis, there are also many examples of metastasis-suppressive lncRNAs in gastrointestinal cancer. Among them, BM742401 exhibits decreased expression in more aggressive cancers, correlating with metastatic properties and decreased survival in gastric cancer tissues [38]. BM742401 over-expression significantly reduced the size and number of foci and inhibited cancer metastasis

by regulating MMP9. PCAT-1 (prostate cancer-associated ncRNA transcripts1), a lncRNA located in the chromosome 8q24 gene desert and -725 kb upstream of the c-MYC oncogene, has been discovered by RNA sequence and implicated in the disease progression of patients with prostate cancer [61]. Recently, Ge et al. [62] found that PCAT-1 was also up-regulated in colon cancer, and there was a significant association between PCAT-1 expression and distant metastasis, but not other clinical characteristics.

LncRNAs and cancer stem cells

The role of lncRNAs in the biology of cancer stem cells begins to gain appreciation in the study of breast cancer, although how lncRNAs affect the progression of gastric and intestinal cancer remains unknown. It has been reported that Abexinostat, a histone deacetylase inhibitor, can induce differentiation of breast cancer stem cells, accompanied with low lncRNA Xist expression [63]. Most recently, lncRNA-ROR has been reported to regulate an epithelial-to-mesenchymal transition (EMT) program in immortalized human mammary epithelial cells and enhanced breast cancer cell migration and invasion as well as generated stem cell properties [64].

Conclusions

LncRNAs have been emerging rapidly as a diverse group of important regulators of genetic information flow that interact with the epigenetic, transcriptional, and posttranscriptional regulation. Functional alterations of specific lncRNAs promote tumor formation and progression in many human malignancies including gastrointestinal cancer. Some lncRNAs may provide new clues for the diagnosis and treatment of gastrointestinal cancer. However, only a small part of gastrointestinal cancer-related lncRNAs have been well characterized. The vast majority of lncRNAs exhibit much lower abundance as compared with typical protein-coding mRNAs, raising a precaution that some of them might be the product of transcriptional noise without any biological function. Moreover, the functional characteristics of individual lncRNAs have often been solely based on their expression pattern correlation with neighboring protein-coding genes without any mechanistic understanding. Therefore, more in-depth work will certainly be necessary to accurately appreciate the nature of the lncRNA regulatory networks and their roles in various biological processes as well as human cancers.

Disclosure of conflict of interest

None.

Address correspondence to: Shuang Wang or Yanqing Ding, Department of Pathology, School of Basic Medical Sciences, Southern Medical University, Guangzhou, 510515, China. E-mail: shuangw@126.com (SW); dyqgz@126.com (YQD)

References

[1] Crick F. Central dogma of molecular biology. Nature 1970; 227: 561-563.

[2] Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermuller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung WK, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei CL, Ruan Y, Struhl K, Gerstein M, Antonarakis SE, Fu Y, Green ED, Karaoz U, Siepel A, Taylor J, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Cooper GM, Asimenos G, Dewey CN, Hou M, Nikolaev S, Montoya-Burgos JI, Loytynoja A, Whelan S, Pardi F, Massingham T, Huang H, Zhang NR,

Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W, Sidow A, Trinklein ND, Zhang ZD, Barrera L, Stuart R, King DC, Ameur A, Enroth S, Bieda MC, Kim J, Bhinge AA, Jiang N, Liu J, Yao F, Vega VB, Lee CW, Ng P, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Fowler JC, Couttet P, Bruce AW, Dovey OM, Ellis PD, Langford CF, Nix DA, Euskirchen G, Hartman S, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Xu M, Haidar JN, Yu Y, Iyer VR, Green RD, Wadelius C, Farnham PJ, Ren B, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Armengol L, Bird CP, de Bakker PI, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Woodroffe A, Davydov E, Dimas A, Eyras E, Hallgrimsdottir IB, Huppert J, Zody MC, Abecasis GR, Estivill X, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VV, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B and de Jong PJ. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007; 447: 799-816.

[3] Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakrabortty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigó R, Gingeras TR. Landscape of transcription in human cells. Nature 2012; 489: 101-108.

[4] Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H and Gingeras TR. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 2007; 316: 1484-1488.

[5] Mattick JS. The genetic signatures of noncoding RNAs. PLoS Genet 2009; 5: e1000459.

[6] Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A and Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011; 25: 1915-1927.

[7] Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL and Lander ES. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009; 458: 223-227.

[8] Amaral PP, Dinger ME, Mercer TR and Mattick JS. The eukaryotic genome as an RNA machine. Science 2008; 319: 1787-1789.

[9] Rinn JL and Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem 2012; 81: 145-166.

[10] Niazi F and Valadkhan S. Computational analysis of functional long noncoding RNAs reveals lack of peptide-coding capacity and parallels with 3' UTRs. RNA 2012; 18: 825-843.

[11] Novikova IV, Hennelly SP and Sanbonmatsu KY. Sizing up long non-coding RNAs: do lncRNAs have secondary and tertiary structure? Bioarchitecture 2012; 2: 189-199.

[12] Halvorsen M, Martin JS, Broadaway S and Laederach A. Disease-associated mutations that alter the RNA structural ensemble. PLoS Genet 2010; 6: e1001074.

[13] Xu C, Yang M, Tian J, Wang X and Li Z. MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. Int J Oncol 2011; 39: 169-175.

[14] Volders PJ, Helsens K, Wang X, Menten B, Martens L, Gevaert K, Vandesompele J and Mestdagh P. LNCipedia: a database for annotated human lncRNA transcript sequences and structures. Nucleic Acids Res 2013; 41: D246-251.

[15] Novikova IV, Hennelly SP and Sanbonmatsu KY. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. Nucleic Acids Res 2012; 40: 5034-5051.

[16] Wilusz JE, Freier SM and Spector DL. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell 2008; 135: 919-932.
[17] Barsotti AM, Beckerman R, Laptenko O, Huppi K, Caplen NJ and Prives C. p53-Dependent induction of PVT1 and miR-1204. J Biol Chem 2012; 287: 2509-2519.
[18] Graham LD, Pedersen SK, Brown GS, Ho T, Kassir Z, Moynihan AT, Vizgoft EK, Dunne R, Pimlott L, Young GP, Lapointe LC and Molloy PL. Colorectal Neoplasia Differentially Expressed (CRNDE), a Novel Gene with Elevated Expression in Colorectal Adenomas and Adenocarcinomas. Genes Cancer 2011; 2: 829-840.
[19] Venables JP, Klinck R, Koh C, Gervais-Bird J, Bramard A, Inkel L, Durand M, Couture S, Froehlich U, Lapointe E, Lucier JF, Thibault P, Rancourt C, Tremblay K, Prinos P, Chabot B and Elela SA. Cancer-associated regulation of alternative splicing. Nat Struct Mol Biol 2009; 16: 670-676.
[20] Ponting CP, Oliver PL and Reik W. Evolution and functions of long noncoding RNAs. Cell 2009; 136: 629-641.
[21] Mercer TR, Dinger ME and Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet 2009; 10: 155-159.
[22] Clark MB, Johnston RL, Inostroza-Ponta M, Fox AH, Fortini E, Moscato P, Dinger ME and Mattick JS. Genome-wide analysis of long noncoding RNA stability. Genome Res 2012; 22: 885-898.
[23] Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES and Rinn JL. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A 2009; 106: 11667-11672.
[24] Brown CJ, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, Lawrence J and Willard HF. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 1992; 71: 527-542.
[25] Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S and Rastan S. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell 1992; 71: 515-526.
[26] Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R and Willard HF. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 1991; 349: 38-44.
[27] Clark MB and Mattick JS. Long noncoding RNAs in cell biology. Semin Cell Dev Biol 2011; 22: 366-376.
[28] Wilusz JE, Sunwoo H and Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 2009; 23: 1494-1504.
[29] Gibb EA, Brown CJ and Lam WL. The functional role of long non-coding RNA in human carcinomas. Mol Cancer 2011; 10: 38.
[30] Huarte M and Rinn JL. Large non-coding RNAs: missing links in cancer? Hum Mol Genet 2010; 19: R152-161.
[31] Song H, Sun W, Ye G, Ding X, Liu Z, Zhang S, Xia T, Xiao B, Xi Y and Guo J. Long non-coding RNA expression profile in human gastric cancer and its clinical significances. J Transl Med 2013; 11: 225.
[32] Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S and Mori M. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res 2011; 71: 6320-6326.
[33] Endo H, Shiroki T, Nakagawa T, Yokoyama M, Tamai K, Yamanami H, Fujiya T, Sato I, Yamaguchi K, Tanaka N, Iijima K, Shimosegawa T, Sugamura K and Satoh K. Enhanced expression of long non-coding RNA HOTAIR is associated with the development of gastric cancer. PLoS One 2013; 8: e77070.
[34] Yang F, Bi J, Xue X, Zheng L, Zhi K, Hua J and Fang G. Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. FEBS J 2012; 279: 3159-3165.
[35] Yang F, Xue X, Bi J, Zheng L, Zhi K, Gu Y and Fang G. Long noncoding RNA CCAT1, which could be activated by c-Myc, promotes the progression of gastric carcinoma. J Cancer Res Clin Oncol 2013; 139: 437-445.
[36] Ling H, Spizzo R, Atlasi Y, Nicoloso M, Shimizu M, Redis RS, Nishida N, Gafa R, Song J, Guo Z, Ivan C, Barbarotto E, De Vries I, Zhang X, Ferracin M, Churchman M, van Galen JF, Beverloo BH, Shariati M, Haderk F, Estecio MR, Garcia-Manero G, Patijn GA, Gotley DC, Bhardwaj V, Shureiqi I, Sen S, Multani AS, Welsh J, Yamamoto K, Taniguchi I, Song MA, Gallinger S, Casey G, Thibodeau SN, Le Marchand L, Tiirikainen M, Mani SA, Zhang W, Davuluri RV, Mimori K, Mori M, Sieuwerts AM, Martens JW, Tomlinson I, Negrini M, Berindan-Neagoe I, Foekens JA, Hamilton SR, Lanza G, Kopetz S, Fodde R and Calin GA. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res 2013; 23: 1446-1461.

[37] Zhai H, Fesler A, Schee K, Fodstad O, Flatmark K and Ju J. Clinical significance of long intergenic noncoding RNA-p21 in colorectal cancer. Clin Colorectal Cancer 2013; 12: 261-266.
[38] Park SM, Park SJ, Kim HJ, Kwon OH, Kang TW, Sohn HA, Kim SK, Moo Noh S, Song KS, Jang SJ, Sung Kim Y and Kim SY. A known expressed sequence tag, BM742401, is a potent lincRNA inhibiting cancer metastasis. Exp Mol Med 2013; 45: e31.
[39] Kam Y, Rubinstein A, Naik S, Djavsarov I, Halle D, Ariel I, Gure AO, Stojadinovic A, Pan H, Tsivin V, Nissan A and Yavin E. Detection of a long non-coding RNA (CCAT1) in living cells and human adenocarcinoma of colon tissues using FIT-PNA molecular beacons. Cancer Lett 2013; [Epub ahead of print].
[40] Sun W, Wu Y, Yu X, Liu Y, Song H, Xia T, Xiao B and Guo J. Decreased expression of long noncoding RNA AC096655.1-002 in gastric cancer and its clinical significance. Tumour Biol 2013; 34: 2697-2701.
[41] Sun M, Xia R, Jin F, Xu T, Liu Z, De W and Liu X. Downregulated long noncoding RNA MEG3 is associated with poor prognosis and promotes cell proliferation in gastric cancer. Tumour Biol 2014; 35: 1065-1073.
[42] Huang X, Yuan T, Tschannen M, Sun Z, Jacob H, Du M, Liang M, Dittmar RL, Liu Y, Kohli M, Thibodeau SN, Boardman L and Wang L. Characterization of human plasma-derived exosomal RNAs by deep sequencing. BMC Genomics 2013; 14: 319.
[43] Ren S, Wang F, Shen J, Sun Y, Xu W, Lu J, Wei M, Xu C, Wu C, Zhang Z, Gao X, Liu Z, Hou J and Huang J. Long non-coding RNA metastasis associated in lung adenocarcinoma transcript 1 derived miniRNA as a novel plasma-based biomarker for diagnosing prostate cancer. Eur J Cancer 2013; 49: 2949-2959.
[44] Arita T, Ichikawa D, Konishi H, Komatsu S, Shiozaki A, Shoda K, Kawaguchi T, Hirajima S, Nagata H, Kubota T, Fujiwara H, Okamoto K and Otsuji E. Circulating long non-coding RNAs in plasma of patients with gastric cancer. Anticancer Res 2013; 33: 3185-3193.
[45] Hanahan D and Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144: 646-674.
[46] Yang MH, Yu J, Chen N, Wang XY, Liu XY, Wang S and Ding YQ. Elevated microRNA-31 expression regulates colorectal cancer progression by repressing its target gene SATB2. PLoS One 2013; 8: e85353.
[47] Yang MH, Yu J, Jiang DM, Li WL, Wang S and Ding YQ. MicroRNA-182 targets special AT-rich sequence-binding protein 2 to promote colorectal cancer proliferation and metastasis. J Transl Med 2014; 12: 109.
[48] Guan Y, Kuo WL, Stilwell JL, Takano H, Lapuk AV, Fridlyand J, Mao JH, Yu M, Miller MA, Santos JL, Kalloger SE, Carlson JW, Ginzinger DG, Celniker SE, Mills GB, Huntsman DG and Gray JW. Amplification of PVT1 contributes to the pathophysiology of ovarian and breast cancer. Clin Cancer Res 2007; 13: 5745-5755.
[49] Takahashi Y, Sawada G, Kurashige J, Uchi R, Matsumura T, Ueo H, Takano Y, Eguchi H, Sudo T, Sugimachi K, Yamamoto H, Doki Y, Mori M and Mimori K. Amplification of PVT-1 is involved in poor prognosis via apoptosis inhibition in colorectal cancers. Br J Cancer 2014; 110: 164-171.
[50] Zhou Y, Zhang X and Klibanski A. MEG3 noncoding RNA: a tumor suppressor. J Mol Endocrinol 2012; 48: R45-53.
[51] Benetatos L, Vartholomatos G and Hatzimichael E. MEG3 imprinted gene contribution in tumorigenesis. Int J Cancer 2011; 129: 773-779.
[52] Yang F, Xue X, Bi J, Zheng L, Zhi K, Gu Y and Fang G. Long noncoding RNA CCAT1, which could be activated by c-Myc, promotes the progression of gastric carcinoma. J Cancer Res Clin Oncol 2013; 139: 437-445.
[53] Yang F, Bi J, Xue X, Zheng L, Zhi K, Hua J and Fang G. Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. FEBS J 2012; 279: 3159-3165.
[54] Zhuang M, Gao W, Xu J, Wang P and Shu Y. The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. Biochem Biophys Res Commun 2014; 448: 315-322.
[55] Du Z, Fei T, Verhaak RG, Su Z, Zhang Y, Brown M, Chen Y and Liu XS. Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. Nat Struct Mol Biol 2013; 20: 908-913.
[56] Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, Tidow N, Brandt B, Buerger H, Bulk E, Thomas M, Berdel WE, Serve H and Muller-Tidow C. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene 2003; 22: 8031-8041.
[57] Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E and Chang HY. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007; 129: 1311-1323.
[58] Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S and Mori M. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associ-

ated with poor prognosis in colorectal cancers. Cancer Res 2011; 71: 6320-6326.

[59] Endo H, Shiroki T, Nakagawa T, Yokoyama M, Tamai K, Yamanami H, Fujiya T, Sato I, Yamaguchi K, Tanaka N, Iijima K, Shimosegawa T, Sugamura K and Satoh K. Enhanced expression of long non-coding RNA HOTAIR is associated with the development of gastric cancer. PLoS One 2013; 8: e77070.

[60] Xu ZY, Yu QM, Du YA, Yang LT, Dong RZ, Huang L, Yu PF and Cheng XD. Knockdown of long non-coding RNA HOTAIR suppresses tumor invasion and reverses epithelial-mesenchymal transition in gastric cancer. Int J Biol Sci 2013; 9: 587-597.

[61] Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, Laxman B, Asangani IA, Grasso CS, Kominsky HD, Cao X, Jing X, Wang X, Siddiqui J, Wei JT, Robinson D, Iyer HK, Palanisamy N, Maher CA and Chinnaiyan AM. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotechnol 2011; 29: 742-749.

[62] Ge X, Chen Y, Liao X, Liu D, Li F, Ruan H and Jia W. Overexpression of long noncoding RNA PCAT-1 is a novel biomarker of poor prognosis in patients with colorectal cancer. Med Oncol 2013; 30: 588.

[63] Salvador MA, Wicinski J, Cabaud O, Toiron Y, Finetti P, Josselin E, Lelièvre H, Kraus-Berthier L, Depil S, Bertucci F, Collette Y, Birnbaum D, Charafe-Jauffret E, Ginestier C. The histone deacetylase inhibitor abexinostat induces cancer stem cells differentiation in breast cancer with low Xist expression. Clin Cancer Res 2013; 19: 6520-31.

[64] Hou P, Zhao Y, Li Z, Yao R, Ma M, Gao Y, Zhao L, Zhang Y, Huang B, Lu J. LincRNA-ROR induces epithelial-to-mesenchymal transition and contributes to breast cancer tumorigenesis and metastasis. Cell Death Dis 2014; 5: e1287.