

# Long Noncoding RNAs in Breast Cancer: Implications for Pathogenesis, Diagnosis, and Therapy

Jaroslav Juráček, Robert Iliev, Marek Svoboda, and Ondrej Slaby

## Abstract

Human genome mapping has revealed that protein-coding genes represent less than 2 % of the total genome sequence, and simultaneously more than 75 % of the genome is actively transcribed into RNA. Recent studies of the human transcriptome led to the discovery of new heterogeneous group of transcripts—noncoding RNAs. The major part of these ncRNAs consists of long noncoding RNAs (lncRNAs), which differ in size, location on genome, and biological functions. Generally, through distinct mechanisms, they affect a number of biological processes, such as modulation of protein activity, alternative splicing of mRNA, and epigenetic regulation or microRNA silencing, and play a key role in transcriptional and post-transcriptional gene expression regulation. Deregulated levels of lncRNAs were observed with a wide range of tumors, including breast cancer. Gene expression patterns of lncRNAs are able to distinguish normal and tumor tissue or even various breast cancer stages, which makes them a potential diagnostic and prognostic biomarkers or therapeutic targets.

## Keywords

Breast cancer • Long noncoding RNA • lncRNA • lincRNA • Diagnosis • Prognosis • Therapeutic targets

## Introduction

Breast cancer is the most frequently diagnosed cancer and also the leading cause of cancer death among females worldwide [1]. It is a heterogeneous disease, which can be divided by various approaches into many molecular subgroups. Despite the notable progress in diagnosis and therapy, a significant number of breast cancer patients with the same diagnostic profile indicate distinctly different clinical outcomes [2, 3].

---

J. Juráček, BS • R. Iliev, MS  
M. Svoboda, PhD, MD • O. Slaby, PhD (✉)  
Research Group Molecular Oncology II, Central  
European Institute of Technology, Masaryk  
University, Brno, Czech Republic  
e-mail: [on.slaby@gmail.com](mailto:on.slaby@gmail.com)

This diversity presents a challenge for better molecular classification and personalized therapy. Therefore, one of the current goals in breast cancer research is to find sensitive and specific noninvasive biomarkers, which can be used for early stage breast cancer detection, as well as for monitoring of the disease and response to therapy [4]. Such biomolecules include long noncoding RNAs, recently described in various types of cancer, including breast cancer.

Despite the central dogma of molecular biology, which understands RNAs as a tool for protein synthesis, there is a large number of RNAs which, instead of coding the protein, act as functional RNA. Moreover, with development of transcriptome analytical methods, it was found that these noncoding RNAs constitute the major part of the human genome [5]. Originally, these RNAs were considered as waste, but now it is clear that they significantly affect diverse cellular pathways. An important and biggest group of noncoding RNAs are the long noncoding RNAs, endogenous cellular molecules with a length of 200 nt to 100 kb [6]. lncRNAs are often capped, polyadenylated, and spliced, yet do not overlap other protein-coding genes [7]. Recent research has shown that lncRNAs plays a key role in transcriptional and posttranscriptional regulation of gene expression [8]. Unsurprisingly, as with other noncoding RNAs, different expressions of lncRNAs were observed in tumor and non-tumor tissue as well. This fact confirms involvement of long noncoding RNAs in tumorigenesis [9].

The regulatory function of RNAs was described in 1961 by François Jacob and Jacques Monod [10], whereas first individual lncRNAs H19 and XIST (X-inactive-specific transcript), which is critical to X chromosome inactivation, were identified a few decades later [11, 12]. However, these RNAs were only discovered to be non-protein coding; therefore, a true milestone was a study by Okazaki et al., who studied mouse genome using large-scale sequencing and defined lncRNAs as a separate class of transcripts [13]. It is estimated that the number of lncRNAs is 7,000–23,000; nevertheless, the list of functionally validated lncRNAs is much shorter (approximately 200lncRNAs) [14].

The aim of this chapter is to provide a short introduction to long noncoding RNAs classification and biology. We also describe basic methods for high-throughput and functional analysis of lncRNAs. The main part is focused on the roles of lncRNAs in pathogenesis and diagnosis of breast cancer and their potential usage in prediction of prognosis or targeted therapy.

---

## Classification and Biology of Long Noncoding RNAs

A growing number of newly annotated long noncoding RNAs together with their varied length and biological functions explain the need for a clear categorization. Some classification approaches, like categorization by position relative to coding genes, splicing, and polyadenylation status or by molecular mechanism [15–17], are not accurate because individual lncRNAs may represent more subgroups. Although we currently lack satisfactory classifications for these transcripts, here we summarize recently discovered groups of lncRNAs as long intergenic noncoding RNAs, long intronic noncoding RNAs, long ncRNAs with dual functions, telomere-associated lncRNAs, pseudogene RNAs, and transcribed-ultraconserved regions [18] (See Table 7.1).

Long noncoding RNAs are a diverse group of transcripts, which differ in size, location in the genome, and other biological properties. Such diversity is also the reason why a wide range of functions was observed in lncRNAs. Long ncRNAs can affect gene expression via RNA polymerase II inhibition (B2 SINE) or chromatin modification (COLDAIR). They can also serve as precursors of siRNAs (H19) and other small ncRNAs (GAS5). By forming complexes with proteins, they may modulate its activity (SRA), influence structural and regulatory functions (XIST), change protein localization, or affect epigenetic processes (HOTAIR). They are also involved in alternative splicing of mRNA (MALAT-1) and are responsible for microRNA silencing (HULC) (See Fig. 7.1).

**Table 7.1** Classification of lincRNAs: their characteristics and meaning in biological processes and diseases

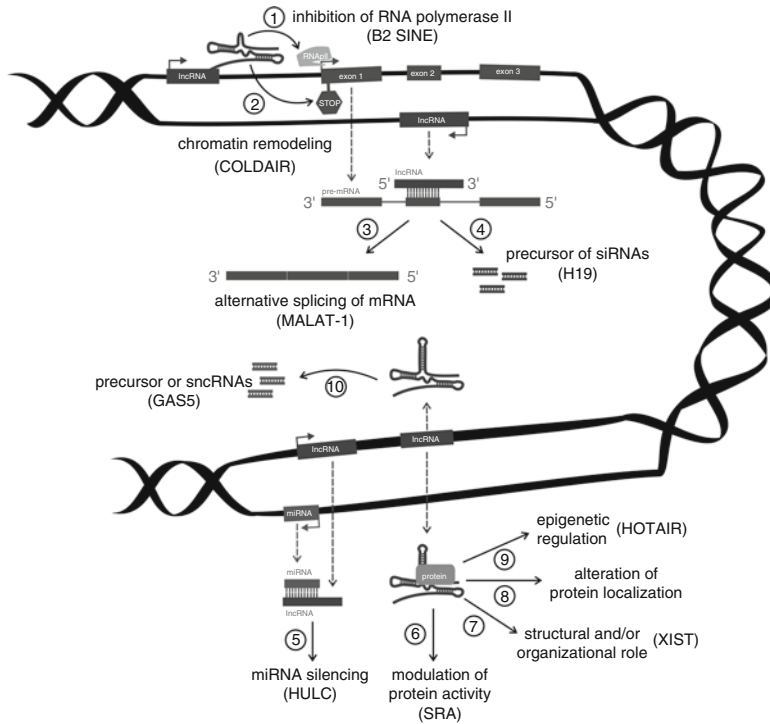
Class	Symbol	Characteristic	Disease/biological function associations
Long intergenic noncoding RNAs	lincRNAs	Ranging from hundreds to tens of thousands nts; lie within the genomic intervals between two genes; transcriptional cis-regulation of neighboring genes	Involved in tumorigenesis and cancer metastasis/involved in diverse biological processes such as dosage compensation and/or imprinting
Long intronic noncoding RNAs		Lie within the introns; evolutionary conserved; tissue and subcellular expression specific	Aberrantly expressed in human cancers/possible link with posttranscriptional gene silencing
Telomere-associated lincRNAs	TERRAs	100 bp–9 kb; conserved among eukaryotes; synthesized from C-rich strand; polyadenylated; form inter-molecular G-quadruplex structure with single-stranded telomeric DNA	Possible impact on telomere-associated diseases including many cancers/negative regulation of telomere length and activity through inhibition of telomerase
Long noncoding RNAs with dual functions		Both protein-coding and functionally regulatory RNA capacity	Deregulation described in breast and ovarian tumors/modulate gene expression through diverse mechanisms
Pseudogene RNAs		Gene copies that have lost the ability to code for a protein; potential to regulate their homologous protein-coding genes; made through retrotransposition; tissue specific	Often deregulated during tumorigenesis and cancer progression/regulation of tumor suppressors and oncogenes by acting as microRNA decoys
Transcribed-ultraconserved regions	T-UCRs	Longer than 200 bp; absolutely conserved between orthologous regions of human, rat, and mouse; located in especially intra- and intergenic regions	Expression is often altered in some cancers; possible involvement in tumorigenesis/antisense inhibitors for protein-coding genes or other ncRNAs

## Long Intergenic Noncoding RNAs

Long intergenic noncoding RNAs (also known as lincRNAs) were originally identified by Guttman et al., who used methods to reconstruct the transcriptome of a mammalian cell [19]. They described an evolutionary conserved group of noncoding RNAs with a range in length from a few hundred to tens of thousands of bases. Genes of lincRNAs are localized in regions of DNA between two protein-coding genes, but they lack any protein-coding capacity and open reading frames. To date, more than 8,000 lincRNAs have been identified, but most of them remain unannotated [8, 20], and therefore the functions of

lincRNAs are largely unknown. However, the involvement of lincRNAs in biological processes has been found, including cell-cycle regulation, imprinting, embryonic stem cell pluripotency, and cell proliferation [21, 22].

The most unclear part of lincRNAs biology is the basis of their molecular mechanism of action. In this context, targeting of chromatin modification complexes (i.e., histone-modifying enzymes) is frequently mentioned, which directly leads to gene expression regulation [8]. Some lincRNAs interact with numerous effector proteins and thus control their levels (i.e., XIST). Others can affect alternative splicing (MALAT-1) by controlling levels of splicing factors [23]. Recently,



**Fig. 7.1** Schematic illustration of lincRNAs functioning. lincRNA transcribed from an upstream noncoding promoter can negatively (1) or positively (2) affect expression of the downstream gene by inhibiting RNA polymerase II recruitment and/or inducing chromatin remodeling, respectively. lincRNA is able to hybridize to the pre-mRNA and block recognition of the splice sites by the spliceosome, thus resulting in an alternatively spliced transcript (3). Alternatively, hybridization of the

sense and antisense transcripts can allow Dicer to generate endogenous siRNAs (4). The binding of lincRNA to the miRNA results in the miRNA function silencing (5). The complex of lincRNA and specific protein partners can modulate the activity of the protein (6), is involved in structural and organization roles of the cell (7), alters the protein localizes in the cell (8), and affects epigenetic processes (9). Finally, long ncRNAs can be processed to the small RNAs (10)

it has been demonstrated that lincRNAs can act as competitive inhibitors of microRNAs termed “microRNA sponges.” Such molecules contain binding sites for specific miRNAs and thus may regulate their level [24].

As mentioned, lincRNAs regulate significant number of genes, but their expression is under genetic control as well [25]. Individual lincRNAs are transcriptionally regulated by important transcription factors such as p53, NFκB, Sox2, Oct4, and Nanog [19]. Interestingly, Juan et al. suggest that some miRNAs may bind lincRNAs and cause their repression [26].

Relevance for translational medicine stems from the fact that long intergenic noncoding RNAs are remarkably tissue specific and deregulated within a large number of diseases,

including cancer. Different levels of lincRNAs were observed at various stages of breast cancer, so based on these gene expression patterns they may serve as potential prognostic markers [9].

## Long Intronic Noncoding RNAs

More than one-third of conserved noncoding regions in human genome consist of intronic regions. John Mattick first suggested that sequence of introns is not random and that introns may be involved in gene regulation [27]. According to this, it was found that about 81 % of human protein-coding genes have transcriptionally active introns [28]. Finally, discovery of numerous evolutionarily conserved regions

in introns that match the size of lncRNAs led to the identification of a new class of transcripts—long intronic noncoding RNAs. These RNAs are exclusively expressed in the nucleus, and it is expected that intronic ncRNA expression would be responsive to common physiological signals, e.g., hormones [29]. The biogenesis is poorly understood, but involvement of RNA polymerase II (RNAP II) is assumed. The presence of poly(A+) tail may serve as indirect evidence [18].

The main role of intronic noncoding RNAs is posttranscriptional regulation of gene expression. In his study, Louro et al. described some mechanisms by which RNAs can regulate gene expression [30]. Interestingly, it was found that intronic ncRNAs can serve as precursors of smaller noncoding RNAs. Another mechanism is a direct interaction with promoters, which decrease the expression of the protein-coding RNA. Intronic ncRNAs can also affect RNA alternative splicing by forming RNA–RNA duplexes. Finally, they are probably able to stabilize protein-coding RNA localized on the same locus [28].

Some oncogenes or tumor suppressor genes have noncoding RNAs transcribed from their introns. This might be one of reasons why altered expression of intronic ncRNAs with various malignancies was detected. Moreover, level of long intronic ncRNAs significantly correlates with different degrees in renal, prostate, and pancreatic carcinoma [31–33].

### Long ncRNAs with Dual Functions

According to the central dogma of molecular biology, RNA was considered as an intermediated molecule required for the formation of protein. After the discovery of noncoding RNAs, which in many aspects resembled the mRNA, but lack protein-coding capacity, it was obvious that RNA can act either as a functional or protein-coding molecule. Hence, it was a big surprise when bifunctional RNAs were found. Such RNAs serve both as intermediate molecules translated into protein and as functional RNA [34].

Functions of RNAs are dependent on their secondary and tertiary structure, so the presence of

isoforms can be important for bifunctional character [35]. Well described is the steroid receptor RNA activator (SRA), whose RNA is a noncoding RNA that coactivates several human hormone receptors like progesterone, estrogen, and androgen. Moreover, isoforms of SRA are also expressed to produce proteins. SRA transcripts have been identified in normal human tissues, and increasingly SRA RNA is expressed in breast and ovarian tumors. Interestingly, higher levels of noncoding isoforms of SRA were observed in tumor tissue [36, 37].

### Telomere-Associated lncRNAs

Telomeres—heterochromatic complexes located on linear chromosome ends—are formed by tandem repeats of the TTAGGG sequence. With each run of cell division, one telomeric hexanucleotide is lost, which finally leads to chromosome destabilization. Therefore, telomeres protect chromosomes from degradation and repair activities [38, 39]. Until recently, it was assumed that telomeres are transcriptionally silent, but a recently discovered group of long noncoding RNAs confirmed that telomeres are transcribed into telomeric repeat-containing RNA (TERRA or Tel RNA). TERRA transcripts range between 100 bp and 9 kb and originate in the subtelomeres of telomeric C-rich strand [40, 41]. Based on RNA-FISH techniques, it was identified that TERRA associates with telomeric chromatin [42]. Other studies have suggested a likely role of TERRA in the enzyme telomerase regulation [43]. Finally, TERRA seems to be involved in negative regulation of telomere length [44].

### Pseudogene RNAs

For a long time, pseudogenes were considered as failed copies of coding genes that had lost the capability to produce proteins. Nevertheless, recent research has revealed the ability of pseudogenes to regulate homologous protein-coding genes [45]. Pseudogenes may arise as a result of simple mutations or be generated by retrotransposition,

during which reverse-transcribed RNAs are integrated into the genomic sequence [46, 47]. Many pseudogenes are transcribed into RNA, which can be later processed into smaller RNAs. Thus, gene expression regulation is based on an RNA interference process. Interestingly, pseudogenes can affect gene expression regulation by acting as miRNA decoys [48]. Some studies provide evidence that pseudogenes (i.e., MYLKPI) are involved in cancerogenesis and suggest them as potential diagnostic and therapeutic targets in cancer [49].

### Transcribed-Ultraconserved Regions

The last newly discovered class of ncRNAs is known as transcribed-ultraconserved regions (T-UCRs). Overall, 481 T-UCRs were annotated—all of them are genomic segments of more than 200 base pairs [50]. They are extremely evolutionarily conserved among mammals [51] and are localized especially in intra- and intergenic regions [52]. The degree of conservation may have a fundamental functional importance for ontogeny and phylogeny of mammals. Untranslated UCRs may serve as distal enhancers [53]; on the contrary transcripts they are involved in gene expression regulation as antisense inhibitors for protein-coding genes. Calin et al. found that the expression of many T-UCRs is altered in some types of cancer, especially in adult chronic lymphocytic leukemias, colorectal and hepatocellular carcinomas, and neuroblastomas [54]. Accordingly, it was found that T-UCRs are often located at fragile regions of chromosomes. Specific transcribed-ultraconserved regions are also associated with prognosis and response to therapy, which makes them promising targets in cancer research [53, 55].

---

### Methods for High-Throughput Analysis of Long Noncoding RNAs in Cancer

The human genome mapping has revealed that more than 90 % of the genome is transcribed. Application of high-throughput techniques highlighted the complexity of mammalian transcriptome

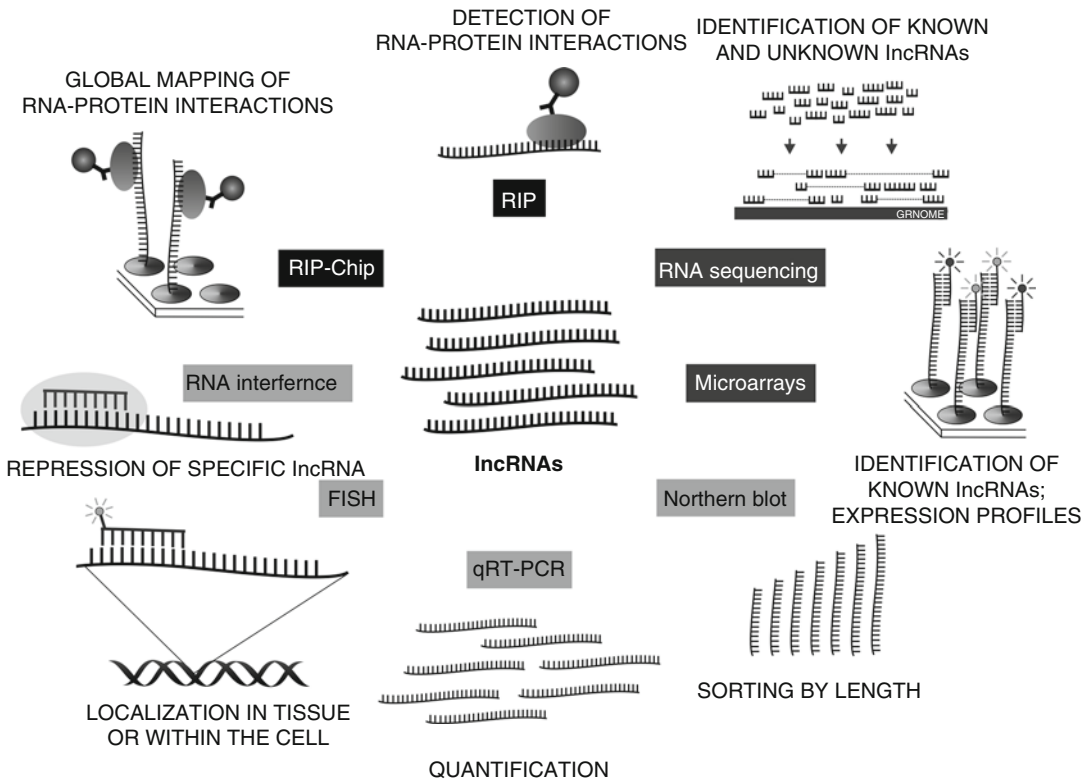
and led to the discovery of long noncoding RNAs, a class of regulatory noncoding RNAs [16, 56]. As a transcriptional class, lncRNAs were first described during the large-scale sequencing of full-length cDNA libraries in the mouse [13]. Such large-scale cDNA analysis and genome annotations can detect or predict thousands of lncRNAs, but their biological functions remain, in most cases, unknown [57]. Methods used in the study of lncRNAs can be thus divided by purpose into (1) high-throughput methods designed for lncRNAs identification (microarrays, RNA sequencing), (2) methods designed for verification of high-throughput data (qRT-PCR, northern blot, FISH, RNAi), and (3) methods designed for detection of RNA–protein interactions (RIP, RIP-CHIP); see Fig. 7.2 [58].

### Microarrays

The structure and expression of long noncoding RNAs is very similar to mRNAs, even though lncRNAs lack open reading frames and other properties necessary for them to be translated into proteins [59]. One of the most common methods of identification is the microarray-based approach. Microarrays are based on nucleic acid hybridization between target molecules and probes and enable simultaneous monitoring of thousands of genes in a single experiment. However, this method shows only whether or not lncRNA is expressed and is therefore not suitable for the identification of novel transcripts [60]. On the other hand, microarrays allow us to detect differences in transcriptional profiles between different tissues and cell types or identify possible targets of lncRNAs [21, 58].

### RNA Sequencing

The transcriptome includes all RNAs synthesized in an organism, including protein-coding, non-coding, alternatively spliced, alternatively polyadenylated, alternatively initiated, sense, antisense, and RNA-edited transcripts [13]. The most widely used method for qualitatively and quantitatively profiling the full set of transcripts is RNA sequencing



**Fig. 7.2** Methods used in the study of lncRNAs can be divided into (1) high-throughput methods designed for lncRNAs identification (microarrays, RNA sequencing),

(2) methods designed for verification of high-throughput data (qRT-PCR, northern blot, FISH, RNAi), (3) methods designed for detection of RNA–protein interactions (RIP, RIP-CHIP)

(RNA-seq), which is based on next-generation sequencing (NGS) [61]. RNA-seq works on a genome-wide scale at single nucleotide resolution and is not limited to detecting already known sequences. Thus, it can be used to discover previously unknown lncRNAs [62]. Considerable disadvantages of this approach are the time and cost related to the downstream analysis of the RNA-seq data [60]. After sequencing, the generated reads are used to assemble the transcriptome, and then novel lncRNAs can be identified and annotated via bioinformatic databases (i.e., FANTOM or ENCODE) [63]. After that, novel lncRNAs often undergo further scrutiny to verify that they are not transcriptional noise and that they indeed do not encode proteins. In the same way, candidate targets are required to be verified by other molecular biology methods as well.

### qRT-PCR and Northern Blot

Quantitative real-time polymerase chain reaction (qRT-PCR) allows amplification and quantification of selected segments of the genome. This highly sensitive method is used for gene expression studies, but qPCR can be used even for analysis of lncRNAs. For this purpose, qPCR is combined with reverse transcription (RT), which ensures transcription of RNA into cDNA. For the verification of high-throughput data, qRT-PCR is frequently followed by northern blot analysis, which is the only direct method to prove the presence of RNA without the need of amplification. This combination was used to demonstrate the lengths of detected lncRNAs and their level of expression [64].

## Fluorescence In Situ Hybridization

Fluorescence in situ hybridization (FISH) is a type of hybridization that uses a fluorescently labeled complementary DNA or RNA strand to localize a specific sequence on a chromosome, section of tissue (in situ) fixed on a slide or even cell [65]. In lncRNAs research, FISH was used to detect individual lncRNA or their localization within the cell [23, 66].

## RNA Interference (RNAi)

The process of RNA interference involves the binding of short interfering RNA molecules to mRNAs, which leads to expression–repression of a gene of interest. The use of synthetic dsRNA allows also effective knockdown of specific lncRNAs and is very important for studying their functions [67].

## RNA Immunoprecipitation and RIP-Chip

Long noncoding RNA may affect the regulation of gene expression also through the modification of chromatin via interactions with various proteins (i.e., transcription factors). Such RNA–protein interactions are easy to detect with RNA immunoprecipitation (RIP). It is an antibody-based technique in which the RNA-binding protein of interest immunoprecipitates together with its associated RNA and allows localize RNA-binding sites on the genome [68]. A number of lncRNAs, such as Xist and Tsix, were identified through this approach [69].

Major advances in the global analysis of subsets of mRNAs bound to RNA-binding proteins brought combinations of RIP and microarrays called RNA-binding protein immunoprecipitation-microarray (Chip) profiling or RIP-Chip [70]. Using this high-throughput method, it was revealed that a large number of lincRNAs associate with chromatin-modifying complexes to affect gene expression [8].

## Long Noncoding RNAs in Pathogenesis of Breast Cancer

lncRNAs were found to be deregulated in several human cancers and show analogically to protein-coding genes tissue-specific expression. Functional studies elucidated a large range of molecular mechanisms used by lncRNAs in cancer cells. Till now, only a few lncRNAs were observed to have altered expression in breast cancer, including HOTAIR, MALAT-1, GAS5, ZFAS1, LSINCT5, SRA1, H19, XIST, and BC200, which are characterized in detail in Table 7.2. Here we discuss molecular functioning of these lncRNAs mainly in the context of typical hallmarks of cancer. Some important molecular mechanisms used by particular lncRNA are mentioned although they are observed in cancer types other than breast cancer.

### HOTAIR

HOX transcript antisense intergenic RNA (HOTAIR) has very important role in cancer metastasis. It was discovered as a 2.2 kb-long ncRNA transcribed in antisense direction from the HOXC gene cluster [71]. HOTAIR functions in trans by interacting and recruiting the polycomb repressive complex 2 (PRC2) to the HOXD locus, which leads to transcriptional silencing across 40 kb. PRC2 complex consisted of H3K27 methylase EZH2, SUZ12, and EED (see Fig. 7.3) [72]. Polycomb group proteins are involved in repression of transcription of large groups of genes. This pathway influences differentiation, pluripotency, and cancer development [73]. Later it was found that HOTAIR interacts with a second histone modification complex, the LSD1/CoREST/REST complex, which coordinates targeting of PRC2 and LSD1 to chromatin for coupled histone H3K27 methylation and K4 demethylation [72]. Overexpression of HOTAIR in epithelial cancer cells alters H3K27 methylation via PRC2 and therefore alters target gene expression. This leads to increased cancer invasiveness and metastases. Therefore, HOTAIR depletion inhibits breast cancer invasiveness [73].

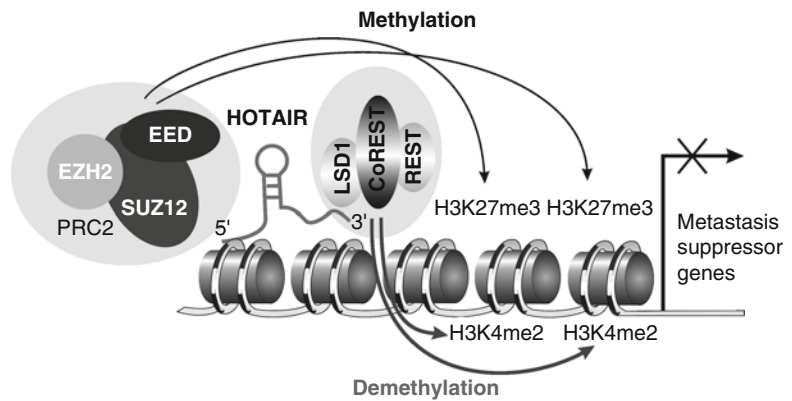


**Table 7.2** Characteristics of lncRNAs deregulated in breast cancer

Gene	Locus	Length (kb)	Function in cancer	Association in hallmarks of cancer	Tumor types	References
HOTAIR	12q13	2.2	Oncogene	Activating invasion and metastasis	BC, CRC, PaC, HCC	[73]
MALAT-1	11q13	8.4	Oncogene	Activating invasion and metastasis; evading growth suppressors	BC, BlaC, HCC	[76, 107, 108]
GAS5	1q25	5.3	Tumor suppressor	Evading growth suppressors	BC, M, Ly, PC	[81, 83, 111]
ZFAS1	20q13	0.5	Tumor suppressor	Evading growth suppressors	BC	[84]
LSINCT5	5p15	2.6	Oncogene	Sustaining proliferative signaling	BC, OC	[85]
SRA1	5q31	0.87	Oncogene	Sustaining proliferative signaling	BC, OC, PC	[90, 92, 93]
XIST	Xq13	42	Oncogene	Activating invasion and metastasis	BC, Ly, OC, CRC	[96, 98]
H19	11p15	3.5	Oncogene, tumor suppressor	Sustaining proliferative signaling	BC, BlaC, CerC, CRC, HCC, LC, M	[100, 101]
BC200	2p21	0.26	Oncogene	Activating invasion and metastasis	BC	[105]

BC breast cancer, CerC cervical cancer, CRC colorectal cancer, PC pancreatic cancer, HCC hepatocellular cancer, BlaC bladder cancer, M melanoma, Ly lymphoma, PC prostate cancer, OV ovarian cancer

**Fig. 7.3** Association of HOTAIR with the polycomb repressive complex 2 (PRC2) and LSD1/CoREST/REST complex



**MALAT-1**

Metastatic-associated lung adenoma transcript 1 (MALAT-1) is abundant in many human cell types. It is probably a very important transcript because its sequence is very conserved across many species. MALAT-1 is an 8,708 nt-long transcript occurring in the nucleus and

frequently localized in nucleus speckles [74]. These structures play a role in pre-mRNA processing. Recently MALAT-1 has been shown to regulate alternative splicing of pre-mRNA by modulating the levels of splicing factors. These factors regulate tissue-specific alternative interactions with SR splicing factor, SRSF1, which affects the distribution of these and other splicing

factors in nuclear speckle domains. Depletion of MALAT-1 with antisense oligonucleotides or transient overexpression of SRSF1 changes the alternative splicing of the endogenous pre-mRNAs. Importantly, MALAT-1 controls cellular phosphorylation status of SR proteins, thereby regulating cellular ratio of phosphorylated versus dephosphorylated form of SR proteins [23], suggesting that MALAT-1 regulates pre-mRNA processing by modulating the levels of active SR proteins. Depletion of MALAT-1 alters the processing of a subset of pre-mRNAs, which play important roles in cancer biology [75].

Recent studies indicate additional functions for MALAT-1 in the nucleus. MALAT-1 was shown to interact with the unmethylated form of CBX4, which controls relocation of growth-control genes between polycomb bodies and interchromatin granules, places of silent or active gene expression, respectively. Altered expression levels of MALAT1 were detected in breast cancer tissue compared to normal breast tissue. Also MALAT-1 locus is frequently altered in breast cancer and other tumor types [76].

## **GAS5**

Growth arrest-specific 5 (GAS5) is the host gene for many snoRNAs, which were found in GAS5 introns, dedicating GAS5 to be involved in the important cellular activities [77]. It was proved that GAS5 transcripts displayed many different patterns of alternate splicing, but there is no putative open reading frame [78]. GAS5 functions as “riborepressor” of the glucocorticoid receptor (GR), influencing cell survival and metabolic activities during starvation by modulating the transcriptional activity of the GR. Its transcript interacts with the DNA-binding domain of GRs and reduces the probability of steroids’ interaction with their receptors [79]. In this way, GAS5 suppresses expression of several genes including cellular inhibitor of apoptosis 2 (cIAP2) and thus sensitizes cells to apoptosis. This induction of apoptosis is independent of other stimuli in several breast cancer cell lines [80]. It was also shown that silencing of endogenous GAS5 levels

in breast cancer cells leads to resistance to apoptosis and various GAS5 transcripts stimulate apoptosis through different cellular signaling pathways [81]. Moreover, in leukemia cell models, GAS5 is required for normal functioning of mammalian target of rapamycin (mTOR) pathway that controls cell growth and proliferation also in breast cancer [82]. In addition, GAS5 locus was found to be frequently altered in many types of cancer (e.g., melanoma, lymphoma, prostate cancer) [83].

## **Zfas1/ZFAS1**

Zinc finger antisense 1 (Zfas1) is a mouse antisense RNA to NFX-1 type containing zinc finger. Zfas1 is located close to a protein-coding gene and in its introns hosts three small nucleolar RNA (snoRNA) genes: Snord12, Snord12b, and Snord12c [84]. Knockdown of Zfas1 in a mammary epithelial cell line resulted in increased cellular proliferation and differentiation. But this knockdown did not substantially alter the levels of the SNORDs. Functional role for Zfas1 in the regulation of alveolar development and epithelial cell differentiation in the mammary gland, together with its dysregulation in human breast cancer, suggests ZFAS1 as a putative tumor suppressor gene. ZFAS1 is highly expressed in the mammary gland and is downregulated in breast tumors compared to normal tissue. While there is relatively low level of primary sequence conservation between Zfas1 and its human ortholog ZFAS1, secondary structures of Zfas1 and ZFAS1 transcripts share several similar features. It was proved that from ZFAS1 mRNA originates at least five different isoforms through alternative splicing.

## **LSINCT5**

Long stress-induced noncoding transcript 5 (LSINCT5) is greatly overexpressed in many of the breast cancer cell lines [85]. LSINCT5 is polyadenylated RNA that is transcribed from the negative strand by RNA polymerase III with no open reading frame. When nuclear and cytoplasmic

LSINCT5 levels are compared, LSINCT5 indicates higher expression in the nuclear fraction. Instead of decrease in cellular proliferation, knockdown of LSINCT5 in cancer-derived cell lines causes expression deregulation of several genes including important kinase (PDPK1), nuclear assembly genes (NEAT1 and PSPC1), genes involved in membrane transport (HERC1), transcription factor (ANKF41), and genes associated with carcinogenesis (EPPK1), cellular stress (PRKAA1/AMPK), motility (ACTR2), and T-cell differentiation (CXCR4, MAPK9/JNK2) [86]. Moreover, LSINCT5 is overexpressed in breast and ovarian cancer tissue.

## SRA1

Steroid receptor RNA activator (SRA) modulates activity of steroid receptors and other transcription factors both at the RNA (SRA) and the protein (SRAP) level [87]. SRA appears highly expressed in the liver, skeletal muscle, adrenal gland, and the pituitary gland, whereas intermediate expression levels are seen in the placenta, lung, kidney, and pancreas. Interestingly, brain and other typical steroid-responsive tissues such as the prostate, breast, uterus, and ovary contained low levels of SRA RNA [88]. SRA is a component of ribonucleoprotein complexes recruited to the promoter of regulated genes. These complexes may contain positive regulators, such as the steroid receptor coactivator 1 (SRC-1), the DExD/H box family of RNA-helicase members p68 and p72, or the pseudouridine synthases Pus1p and Pus3p. Negative regulators, such as SMRT/HDAC1-associated repressor protein (SHARP) and the recently identified SRA stem-loop interacting RNA-binding protein (SLIRP), can also interact with SRA to decrease its activity [89]. Elevated levels of SRA are found in breast tumors and the increased SRA levels might contribute to the altered ER/PR action that occurs during breast tumorigenesis. The SRA1 gene might not only act as an ncRNA but also codes a protein that acts as a coactivator or corepressor. The ratio between noncoding and coding transcripts of SRA1 characterizes specific tumor phenotypes but might also be involved in

breast tumorigenesis and tumor progression by regulating the expression of specific sets of genes [90]. The sequence of the protein encoded by SRA, referred to as SRAP, is highly conserved in Chordata. The most conserved amino acids define two distinct domains (N- and C-terminal) that represent the typical signature of this new family of proteins, and which are likely both participating in SRAP function [91]. This protein is also ubiquitously found in human cancer cell lines derived from the breast [92], the prostate [93], and other tissues, even though levels of expression appear to vary from one cell type to another.

## XIST

X-inactive specific transcript (XIST) is transcribed from the inactivated X chromosome, is involved in its inactivation, and exists in many types of isoforms [94]. On the active X allele, XIST is repressed by its antisense RNA, TSIX [95]. XIST contains a double-hairpin RNA motif in the RepA domain. It is located in the first exon, which is crucial for its ability to bind polycomb repressive complex 2 (PRC2) and propagate epigenetic silencing of the X chromosome [69]. Subsequently, the inactive X (Xi) acquires the typical features of heterochromatin: late replication, hypoacetylation of histones H3 and H4, methylation of histone H3 lysines 9 and 27, lack of methylation of H3 lysine 4, and methylation of DNA CpG islands. Initial studies suggested a role for XIST in hereditary BRCA1-deficient breast cancers [96], whereas data indicated that BRCA1 was not required for XIST to function in these cells [97]. Aberrant XIST regulation was also observed in other cancers, including lymphoma and male testicular germ-cell tumors [98].

## H19

H19 is located in a cluster of imprinted genes on the human chromosome 11. The regulation of H19 is related to its closely linked and reciprocally imprinted neighbor IGF2. They are studied intensively both because of their role in human diseases

and as a model for understanding imprinting control mechanisms. Thereby H19 is transcribed only from the maternal allele, whereas IGF2 expression is exclusively paternal. H19 is considered a regulatory RNA [12]. It has been suggested that H19 functions in many different processes, ranging from transcriptional and posttranscriptional regulation of expression [99] to tumor suppression and oncogenesis, including breast cancer [100]. The expression of H19 is high during vertebrate embryo development, but is downregulated in most tissues shortly after birth with the exception of skeletal tissue and cartilage [101]. In breast cancer cell lines, c-Myc induces the expression of the H19 ncRNA and binds directly to DNA sequence elements called E-boxes close to the imprinting control region (ICR). Thus, c-Myc specifically binds and regulates the active maternal H19 allele and does not bind or affect the expression of the silenced paternal allele. In addition, c-Myc downregulates transcription of the reciprocally imprinted gene IGF2 [103]. H19 was also shown to be directly activated by the oncogenic transcription factor c-Myc in colon cancer, suggesting H19 may be an intermediate functionary between c-Myc and downstream gene expression [102]. The upregulation of H19 by c-Myc and correlation of c-Myc and H19 levels were observed in primary and established tumor cells derived from breast cancer patients [103]. The tumor suppressor gene p53 has been shown to decrease H19 levels. H19 transcripts also serve as a precursor for miR-675, the miRNA involved in the regulation of developmental genes. MiR-675 is processed from the first exon of H19 and leads to a decrease in the levels of tumor suppressor retinoblastoma gene 1 (RB1) [104].

## BC200

BC200 RNA is a 200 nt-long RNA that is selectively expressed in the primate nervous system where it has been identified in somatodendritic domains of a subset of neurons [105]. BC200 RNA is not normally expressed in nonneuronal somatic cells [110]. It has been shown that it is expressed in germ cells and in cultured immortal

cell lines of various nonneural cell types. In order to investigate whether the neuron-specific expression of BC200 RNA is also deregulated during tumorigenesis in nonneural human tissues, 80 different tumor specimens, representing 19 different tumor types, were screened for the presence of this RNA [106]. BC200 was detected in carcinomas of the breast, cervix, esophagus, lung, ovary, parotid, and tongue, but not in corresponding normal tissue. BC200 was not detectable in bladder, colon, kidney, or liver carcinoma tissues examined in this study. These results demonstrate that BC200 expression is deregulated under certain neoplastic conditions. The expression of BC200 RNA in nonneural tumors may indicate a functional interrelationship with induction and progression of these tumors [106].

---

## Long Noncoding RNAs in Diagnosis of Breast Cancer

HOTAIR levels were up to a 2,000-fold increase in primary and metastatic breast cancer tissue compared with normal breast tissue [73]. In breast cancer tissue, moderate or high levels of MALAT-1 were also observed [76]. MALAT-1 levels were increased also in bladder cancer and hepatocellular carcinoma [107, 108]. LSINCT5 had increased expression in breast and ovarian cancer cell line and tumor tissue [85]. Elevated levels of SRA are found in breast tumors and the increased SRA levels might contribute to the altered ER/PR action that occurs during breast tumorigenesis. Relative expression of SRA varies between breast cancer cell lines with different phenotypes [91].

Another type of oncogenic lncRNA is XIST, which is typically expressed by all female somatic cells. But XIST expression has been found to be lost in female breast, ovarian, and cervical cancer cell lines [97]. There is a substantial body of evidence to suggest the occurrence of X chromosome inactivation alterations in breast cancer cells. Interestingly, in cell lines derived from the duct carcinoma of the kidney, XIST gene, along with several other chromosome X genes, was found being amplified [109]. BC200 was expressed in carcinomas of the breast, cervix,

esophagus, lung, ovary, parotid, and tongue but not in corresponding normal tissues [106]. In an independent study, it was shown that BC200 RNA is detectable at significant levels in a number of human tumors, including infiltrating ductal carcinoma of the breast, squamous cell carcinoma of the lung, and lung metastasis of melanoma. Corresponding normal tissue obtained from the same patient was found to be BC200 negative. BC200 RNA is expressed at high levels in high-grade ductal carcinoma in situ (HG DCIS) but not in non-high-grade ductal carcinoma in situ (NHG DCIS) [105]. High expression of BC200 RNA in carcinoma in situ is thus indicative of high grade.

H19 is upregulated by c-Myc and H19 levels were observed in primary and established tumor cells derived from breast cancer patients [103]. In comparison with normal breast epithelial tissue, reduced levels of GAS5 were detected in cancer tissue [80]. ZFAS1 is highly expressed in the mammary gland and is downregulated in breast tumors compared to normal tissue [84]. ZFAS1 expression is decreased in ductal carcinoma relative to normal epithelial cells.

lncRNAs with diagnostic potential in breast cancer are summarized in Table 7.2. The majority of lncRNAs that were identified as deregulated in breast cancer are oncogenes, and their levels in cancers are increased: HOTAIR, MALAT-1, LSINCT5, SRA1, XIST, BC200, and H19. Only two studied lncRNAs indicate properties of tumor suppressors: GAS5 and ZFAS1. At the moment, independent studies in large cohorts of breast cancer patients enabling detailed clinicopathologic correlations are needed to prove and define potential diagnostic usage of these promising lncRNAs.

---

### Long Noncoding RNAs as Prognostic and Predictive Biomarkers in Breast Cancer

From the lncRNAs studied in breast cancer only five indicate the potential to be prognostic biomarkers. In human breast cancer, HOTAIR expression is increased in primary tumors and

metastases, and its expression level in primary tumors positively correlates with the development of metastasis and poor outcomes [72, 73]. Also, MALAT-1 expression was remarkably increased in primary tumors that subsequently metastasized in contrast to primary tumors of patients with better outcomes [107].

Invasive breast cancer cell lines were shown to have higher levels of noncoding SRA than less invasive ones. This suggests that the expression of noncoding SRA in breast cells is probably associated with the ability for invasion [91]. Interestingly, the apparent overexpression of SRAP in some cases correlated with an overall survival in the breast cancer patients [93]. This also suggests that an increase in SRAP expression might characterize a less aggressive phenotype and it is possible that this protein contributes to the improved outcome after tamoxifen antiestrogen therapy [93].

Several studies noted that aggressive breast tumors do not show a detectable Barr body due to the cytological examination of the Xi [111]. Decreased levels of XIST lead to reduced sensitivity to treatment with Taxol in ovarian cancer cell lines, suggesting that the expression of XIST may serve as a predictive biomarker of therapy response [112].

Only the level of GAS5 expression in the breast cancer cell lines showed a general inverse correlation with tumorigenic behavior [82]. Interestingly, in head and neck squamous cell carcinoma, a high level of GAS5 was associated with a good prognosis [113].

Recent evidence suggests that some lncRNAs deregulated in breast cancer tissue may serve as prognostic or predictive biomarkers in breast cancer patients, indicating their potential in translational oncology.

---

### Long Noncoding RNAs as Potential Therapeutic Targets

Besides the imminent use of our knowledge of cancer-associated long ncRNAs for diagnosis, therapeutic applications may be possible in a more distant future. The use of lncRNAs

as therapeutic agents is only beginning to be explored. Although our understanding of the molecular mechanisms of lncRNA function is limited, some features of lncRNAs make them ideal candidates for therapeutic intervention. Many lncRNAs appear to have protein-binding or functional potential that is dependent on secondary structure; this may provide a means of intervention [114]. Preventing the interactions of HOTAIR with the PRC2 or LSD1 complexes, for example, may limit the metastatic potential of breast cancer cells [73]. The progress in the use of RNAi-mediated gene silencing for the treatment of different diseases is encouraging and could be applied to selectively silence oncogenic lncRNAs. Gene therapy could also be applied for the delivery to specific cells of tumor suppressor lncRNAs for the treatment of breast cancer. However, many technical challenges have to be overcome for a wider use of therapeutic RNAi and gene therapy [114]. The expression of the lncRNA H19 is increased in a wide range of human cancers, including breast cancer [100]. One promising therapeutic approach presents a plasmid vector carrying the gene for the diphtheria toxin, which is under the control of the H19 promoter. Intratumoral injections of this plasmid induced high expression levels of diphtheria toxin specifically in the tumor cells, resulting in a reduction of tumor size in human trials [115]. GAS5 expression induces growth arrest and apoptosis independently of other stimuli in some breast cancer cell lines [82]. Therefore, development of technology inducing GAS5 expression in tumors—or designing a vector that would induce the expression of GAS5 when delivered into the tumor cells—might provide an attractive therapeutic approach. Collectively, these advances indicate the significant potential in developing of lncRNAs-mediated therapies.

## Conclusion and Future Perspective

There is much research still on the way toward a deeper understanding of regulation processes, in which lncRNA is one of the important players. Although long noncoding RNA expression

profiles in tumor tissue highlighted the potential value of this class of noncoding RNAs as tumor biomarkers in diagnosis and prognosis of breast cancer patients, only studying the mechanisms of lncRNA involvement in oncogenic and tumor suppressive pathways could lead to the establishment of new diagnostic biomarkers and figure their potential usage as novel therapeutic targets. As the catalog of lncRNAs still grows, it will also become important to elucidate the genetic networks and signaling pathways regulated by the lncRNAs, which are abnormally expressed in breast cancer cells, to understand the role of these lncRNAs in the processes of malignant transformation.

## References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011;61(2):69–90.
2. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A*. 2003;100(18):10393–8.
3. Weigelt B, Reis-Filho JS. Histological and molecular types of breast cancer: is there a unifying taxonomy? *Nat Rev Clin Oncol*. 2009;6(12):718–30.
4. Heneghan HM, Miller N, Lowery AJ, Sweeney KJ, Newell J, Kerin MJ. Circulating microRNAs as novel minimally invasive biomarkers for breast cancer. *Ann Surg*. 2010;251(3):499.
5. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, et al. Global identification of human transcribed sequences with genome tiling arrays. *Science*. 2004;306(5705):2242–6.
6. Chen LL, Carmichael GG. Long noncoding RNAs in mammalian cells: what, where and why? *Wiley Interdiscip Rev RNA*. 2010;1:2–21.
7. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*. 2011;147(7):1537–50.
8. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A*. 2009;106:11667–72.
9. Gutschner T, Diederichs S. The hallmarks of cancer: a long non-coding RNA point of view. *RNA Biol*. 2012;9(6):703–19.
10. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*. 1961;3:318–56.

11. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell*. 2009;136:629–41.
12. Brannan CI, Dees EC, Ingram RS, Tilghman SM. The product of the H19 gene may function as an RNA. *Mol Cell Biol*. 1990;10(1):28–36.
13. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*. 2002;420:563–73.
14. Niazi F, Valadkhan S. Computational analysis of functional long noncoding RNAs reveals lack of peptide-coding capacity and parallels with 3' UTRs. *RNA*. 2012;18:825–43.
15. Mercer TR, Dinger ME, Mattick JS. Long noncoding RNAs: insights into functions. *Nat Rev Genet*. 2009;10:155–9.
16. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell*. 2011;43(6):904–14.
17. Michelhaugh SK, Lipovich L, Blythe J, Jia H, Kapatos G, Bannon MJ. Mining Affymetrix microarray data for long non-coding RNAs: altered expression in the nucleus accumbens of heroin abusers. *J Neurochem*. 2011;116(3):459–66.
18. Sana J, Faltejskova P, Svoboda M, Slaby O. Novel classes of non-coding RNAs and cancer. *J Transl Med*. 2012;10(1):103.
19. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009;458(7235):223–7.
20. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011;25(18):1915–27.
21. Loewer S, Cabili MN, Guttman M, Loh YH, Thomas K, Park IH, et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat Genet*. 2010;42(12):1113–7.
22. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*. 2011;477(7364):295–300.
23. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010;39(6):925–38.
24. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*. 2011;147(2):358–69.
25. Kumar V, Westra HJ, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, et al. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. *PLoS Genet*. 2013;9(1):e1003201.
26. Juan L, Wang G, Radovich M, Schneider BP, Clare SE, Wang Y, et al. Potential roles of microRNAs in regulating long intergenic noncoding RNAs. *BMC Med Genet*. 2013;6 Suppl 1:S7.
27. Mattick JS. Introns: evolution and function. *Curr Opin Genet Dev*. 1994;4(6):823–31.
28. Louro R, Smirnova AS, Verjovski-Almeida S. Long intronic noncoding RNA transcription: expression noise or expression choice? *Genomics*. 2009;93(4):291–8.
29. Louro R, Nakaya HI, Amaral PP, Festa F, Sogayar MC, Da Silva AM, et al. Androgen responsive intronic non-coding RNAs. *BMC Biol*. 2007;5(1):4.
30. Louro R, El-Jundi T, Nakaya HI, Reis EM, Verjovski-Almeida S. Conserved tissue expression signatures of intronic noncoding RNAs transcribed from human and mouse loci. *Genomics*. 2008;92(1):18.
31. Brito GC, Fachel AA, Vettore AL, Vignal GM, Gimber ER, Campos FS, et al. Identification of protein-coding and intronic noncoding RNAs down-regulated in clear cell renal carcinoma. *Mol Carcinog*. 2008;47(10):757–67.
32. Reis EM, Nakaya HI, Louro R, Canavez FC, Flatschart AV, Almeida GT, et al. Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer. *Oncogene*. 2004;23(39):6684–92.
33. Tahira AC, Kubrusly MS, Faria MF, Dazzani B, Fonseca RS, Maracaja-Coutinho V, et al. Long non-coding intronic RNAs are differentially expressed in primary and metastatic pancreatic cancer. *Mol Cancer*. 2011;10:141.
34. Ulveling D, Francastel C, Hubé F. When one is better than two: RNA with dual functions. *Biochimie*. 2011;93(4):633–44.
35. Novikova IV, Hennelly SP, Sanbonmats KY. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res*. 2012;40(11):5034–51.
36. Kawashima H, Takano H, Sugita S, Takahara Y, Sugimura K, Nakatani T. A novel steroid receptor co-activator protein (SRAP) as an alternative form of steroid receptor RNA-activator gene: expression in prostate cancer cells and enhancement of androgen receptor activity. *Biochem J*. 2003;369(Pt 1):163–71.
37. Lanz RB, McKenna NJ, Onate SA, Albrecht U, Wong J, Tsai SY, et al. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell*. 1999;97(1):17–27.
38. Martínez P, Blasco MA. Telomeric and extra-telomeric roles for telomerase and the telomere-binding proteins. *Nat Rev Cancer*. 2011;11(3):161–76.
39. Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E, Lingner J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science*. 2007;318(5851):798–801.
40. Sampl S, Pramas S, Stern C, Preusser M, Marosi C, Holzmann K. Expression of telomeres in astrocytoma WHO grade 2 to 4: TERRA level correlates with telomere length, telomerase activity, and advanced clinical grade. *Transl Oncol*. 2012;5(1):56–65.
41. Schoeftner S, Blasco MA. A 'higher order' of telomere regulation: telomere heterochromatin and telomeric RNAs. *EMBO J*. 2009;28(16):2323–36.

42. Schoeftner S, Blasco MA. Chromatin regulation and non-coding RNAs at mammalian telomeres. *Semin Cell Dev Biol.* 2010;21(2):186–93.
43. Isken O, Maquat LE. Telomeric RNAs as a novel player in telomeric integrity. *F1000 Biol Rep.* 2009;1:90.
44. Luke B, Lingner J. TERRA: telomeric repeat-containing RNA. *EMBO J.* 2009;28(17):2503–10.
45. Harrison PM, Zheng D, Zhang Z, Carriero N, Gerstein M. Transcribed processed pseudogenes in the human genome: an intermediate form of expressed retrosequence lacking protein-coding ability. *Nucleic Acids Res.* 2005;33(8):2374–83.
46. Terai G, Yoshizawa A, Okida H, Asai K, Mituyama T. Discovery of short pseudogenes derived from messenger RNAs. *Nucleic Acids Res.* 2010;38(4):1163–71.
47. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet.* 2000;24(4):363–7.
48. Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DRF. Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA.* 2011;17(5):792–8.
49. Han YJ, Ma SF, Yourek G, Park YD, Garcia JG. A transcribed pseudogene of MYLK promotes cell proliferation. *FASEB J.* 2011;25(7):2305–12.
50. Hudson RS, Yi M, Volfovsky N, Prueitt RL, Esposito D, Volinia S, et al. Transcription signatures encoded by ultraconserved genomic regions in human prostate cancer. *Mol Cancer.* 2013;12(1):13.
51. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, et al. Ultraconserved elements in the human genome. *Science.* 2004;304(5675):1321–5.
52. Mestdagh P, Fredlund E, Pattyn F, Rihani A, Van Maerken T, Vermeulen J, et al. An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastomas. *Oncogene.* 2010;29(24):3583–92.
53. Scaruffi P. The transcribed-ultraconserved regions: a novel class of long noncoding RNAs involved in cancer susceptibility. *ScientificWorldJournal.* 2011;11:340–52.
54. Calin GA, Liu CG, Ferracin M, Hyslop T, Spizzo R, Sevignani C, et al. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell.* 2007;2(3):215–29.
55. Scaruffi P, Stigliani S, Moretti S, Coco S, De Vecchi C, Valdora F, et al. Transcribed-ultra conserved region expression is associated with outcome in high-risk neuroblastoma. *BMC Cancer.* 2009;9:441.
56. Enfield KS, Pikor LA, Martinez VD, Lam WL. Mechanistic roles of noncoding RNAs in lung cancer biology and their clinical implications. *Genet Res Int.* 2012;2012:737416.
57. Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, et al. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science.* 2005;309(5740):1570–3.
58. Yan B, Wang ZH, Guo JT. The research strategies for probing the function of long noncoding RNAs. *Genomics.* 2012;99(2):76–80.
59. Baker M. Long noncoding RNAs: the search for function. *Nat Methods.* 2011;8(5):379.
60. Lee C, Kikyo N. Strategies to identify long noncoding RNAs involved in gene regulation. *Cell Biosci.* 2012;2(1):37.
61. Xuan J, Yu Y, Qing T, Guo L, Shi L. Next-generation sequencing in the clinic: promises and challenges. *Cancer Lett.* 2012;340(2):284–95. pii: S0304-3835(12)00672-6.
62. Atkinson SR, Marguerat S, Bähler J. Exploring long non-coding RNAs through sequencing. *Semin Cell Dev Biol.* 2012;23(2):200–5.
63. Amaral PP, Clark MB, Gascoigne DK, Dinger ME, Mattick JS. lncRNADB: a reference database for long noncoding RNAs. *Nucleic Acids Res.* 2011;39(Database issue):D146–51.
64. Perez DS, Hoage TR, Pritchett JR, Ducharme-Smith AL, Halling ML, Ganapathiraju SC, et al. Long, abundantly expressed non-coding transcripts are altered in cancer. *Hum Mol Genet.* 2008;17(5):642–55.
65. Jehan Z, Uddin S, Al-Kuraya KS. In-situ hybridization as a molecular tool in cancer diagnosis and treatment. *Curr Med Chem.* 2012;19(22):3730–8.
66. Chureau C, Chantalat S, Romito A, Galvani A, Duret L, Avner P, Rougeulle C. Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Hum Mol Genet.* 2011;20(4):705–18.
67. Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science.* 2011;331(6013):76–9.
68. Percipalle P, Obrdlik A. Analysis of nascent RNA transcripts by chromatin RNA immunoprecipitation. *Methods Mol Biol.* 2009;567:215–35.
69. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science.* 2008;322(5902):750–6.
70. Jain R, Devine T, George AD, Chittur SV, Baroni TE, Penalva LO, et al. RIP-chip analysis: RNA-binding protein immunoprecipitation-microarray (chip) profiling. *Methods Mol Biol.* 2011;703:247–63.
71. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* 2007;129(7):1311–23.
72. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science.* 2010;329(5992):689–93.
73. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature.* 2010;464(7291):1071–6.
74. Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics.* 2007;8:39.
75. Lin R, Roychowdhury-Saha M, Black C, Watt AT, Marcusson EG, Freier SM, et al. Control of RNA processing by a large non-coding RNA over-expressed in carcinomas. *FEBS Lett.* 2011;585(4):671–6.



76. Lin R, Maeda S, Liu C, Karin M, Edgington TS. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene*. 2007;26(6):851–8.
77. Smith CM, Steitz JA. Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol Cell Biol*. 1998;18(12):6897–909.
78. Muller AJ, Chatterjee S, Teresky A, Levine AJ. The gas5 gene is disrupted by a frameshift mutation within its longest open reading frame in several inbred mouse strains and maps to murine chromosome 1. *Mamm Genome*. 1998;9(9):773–4.
79. Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal*. 2010;3(107):ra8.
80. Mourtada-Maarabouni M, Hedge VL, Kirkham L, Farzaneh F, Williams GT. Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (GAS5). *J Cell Sci*. 2008;121(Pt 7):939–46.
81. Mourtada-Maarabouni M, Pickard MR, Hedge VL, Farzaneh F, Williams GT. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene*. 2009;28(2):195–208.
82. Mourtada-Maarabouni M, Hasan AM, Farzaneh F, Williams GT. Inhibition of human T-cell proliferation by mammalian target of rapamycin (mTOR) antagonists requires noncoding RNA growth-arrest-specific transcript 5 (GAS5). *Mol Pharmacol*. 2010;78(1):19–28.
83. Nupponen NN, Carpten JD. Prostate cancer susceptibility genes: many studies, many results, no answers. *Cancer Metastasis Rev*. 2001;20(3–4):155–64.
84. Askarian-Amiri ME, Crawford J, French JD, Smart CE, Smith MA, Clark MB, et al. SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. *RNA*. 2011;17(5):878–91.
85. Silva JM, Boczek NJ, Berres MW, Ma X, Smith DI. LSINCT5 is over expressed in breast and ovarian cancer and affects cellular proliferation. *RNA Biol*. 2011;8(3):496–505.
86. Brand S, Dambacher J, Beigel F, Olszak T, Diebold J, Otte JM, et al. CXCR4 and CXCL12 are inversely expressed in colorectal cancer cells and modulate cancer cell migration, invasion and MMP-9 activation. *Exp Cell Res*. 2005;310(1):117–30.
87. Cooper C, Guo J, Yan Y, Chooniedass-Kothari S, Hube F, Hamedani MK, et al. Increasing the relative expression of endogenous non-coding Steroid Receptor RNA Activator (SRA) in human breast cancer cells using modified oligonucleotides. *Nucleic Acids Res*. 2009;37(13):4518–31.
88. Lanz RB, Chua SS, Barron N, Söder BM, DeMayo F, O'Malley BW. Steroid receptor RNA activator stimulates proliferation as well as apoptosis in vivo. *Mol Cell Biol*. 2003;23(20):7163–76.
89. Shi Y, Downes M, Xie W, Kao HY, Ordentlich P, Tsai CC, et al. Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev*. 2001;15(9):1140–51.
90. Rae JM, Johnson MD. What does an orphan G-protein-coupled receptor have to do with estrogen? *Breast Cancer Res*. 2005;7(6):243–4.
91. Leygue E. Steroid receptor RNA activator (SRA1): unusual bifaceted gene products with suspected relevance to breast cancer. *Nucl Recept Signal*. 2007;5:e006.
92. Emberley E, Huang GJ, Hamedani MK, Czosnek A, Ali D, Grolla A, et al. Identification of new human coding steroid receptor RNA activator isoforms. *Biochem Biophys Res Commun*. 2003;301(2):509–15.
93. Kurisu T, Tanaka T, Ishii J, Matsumura K, Sugimura K, Nakatani T, et al. Expression and function of human steroid receptor RNA activator in prostate cancer cells: role of endogenous hSRA protein in androgen receptor-mediated transcription. *Prostate Cancer Prostatic Dis*. 2006;9(2):173–8.
94. Chow J, Heard E. X inactivation and the complexities of silencing a sex chromosome. *Curr Opin Cell Biol*. 2009;21(3):359–66.
95. Lee JT. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev*. 2009;23(16):1831–42.
96. Richardson AL, Wang ZC, De Nicolo A, Lu X, Brown M, Miron A, et al. X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell*. 2006;9(2):121–32.
97. Xiao C, Sharp JA, Kawahara M, Davalos AR, Difilippantonio MJ, Hu Y, et al. The XIST noncoding RNA functions independently of BRCA1 in X inactivation. *Cell*. 2007;128(5):977–89.
98. Kawakami T, Zhang C, Taniguchi T, Kim CJ, Okada Y, Sugihara H, et al. Characterization of loss-of-inactive X in Klinefelter syndrome and female-derived cancer cells. *Oncogene*. 2004;23(36):6163–9.
99. Li YM, Franklin G, Cui HM, Svensson K, He XB, Adam G, et al. The H19 transcript is associated with polysomes and may regulate IGF2 expression in trans. *J Biol Chem*. 1998;273(43):28247–52.
100. Lottin S, Adriaenssens E, Berteaux N, Leprêtre A, Vilain MO, Denhez E, et al. The human H19 gene is frequently overexpressed in myometrium and stroma during pathological endometrial proliferative events. *Eur J Cancer*. 2005;41(1):168–77.
101. Gibb EA, Vucic EA, Enfield KS, Stewart GL, Lonergan KM, Kennett JY, et al. Human cancer long non-coding RNA transcriptomes. *PLoS One*. 2011;6(10):e25915.
102. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*. 2004;10(12):1957–66.
103. Berteaux N, Lottin S, Monté D, Pinte S, Quatannens B, Coll J, et al. H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J Biol Chem*. 2005;280(33):29625–36.

104. Tsang WP, Ng EK, Ng SS, Jin H, Yu J, Sung JJ, et al. Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis*. 2010;31(3):350–8.
105. Iacoangeli A, Lin Y, Morley EJ, Muslimov IA, Bianchi R, Reilly J, et al. BC200 RNA in invasive and preinvasive breast cancer. *Carcinogenesis*. 2004;25(11):2125–33.
106. Chen W, Böcker W, Brosius J, Tiedge H. Expression of neural BC200 RNA in human tumours. *J Pathol*. 1997;183(3):345–51.
107. Ying L, Chen Q, Wang Y, Zhou Z, Huang Y, Qiu F. Upregulated MALAT-1 contributes to bladder cancer cell migration by inducing epithelial-to-mesenchymal transition. *Mol Biosyst*. 2012;8(9):2289–94.
108. Lai MC, Yang Z, Zhou L, Zhu QQ, Xie HY, Zhang F, et al. Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Med Oncol*. 2012;29(3):1810–6.
109. Wu ZS, Lee JH, Kwon JA, Kim SH, Han SH, An JS, et al. Genetic alterations and chemosensitivity profile in newly established human renal collecting duct carcinoma cell lines. *BJU Int*. 2009;103(12):1721–8.
110. Tiedge H, Chen W, Brosius J. Primary structure, neural-specific expression, and dendritic location of human BC200 RNA. *J Neurosci*. 1993;13(6):2382–90.
111. Huang KC, Rao PH, Lau CC, Heard E, Ng SK, Brown C, et al. Relationship of XIST expression and responses of ovarian cancer to chemotherapy. *Mol Cancer Ther*. 2002;1(10):769–76.
112. Pageau GJ, Hall LL, Ganesan S, Livingston DM, Lawrence JB. The disappearing Barr body in breast and ovarian cancers. *Nat Rev Cancer*. 2007;7(8):628–33.
113. Gee HE, Buffa FM, Camps C, Ramachandran A, Leek R, Taylor M, et al. The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. *Br J Cancer*. 2011;104(7):1168–77.
114. Gibb EA, Brown CJ, Lam WL. The functional role of long non-coding RNA in human carcinomas. *Mol Cancer*. 2011;10:38.
115. Amit D, Hochberg A. Development of targeted therapy for a broad spectrum of cancers (pancreatic cancer, ovarian cancer, glioblastoma and HCC) mediated by a double promoter plasmid expressing diphtheria toxin under the control of H19 and IGF2-P4 regulatory sequences. *Int J Clin Exp Med*. 2012;5(4):296–305.