JYI Journal of Young Investigators RESEARCH REVIEW

The RNA HOTAIR Promotes Chromatin Alteration in Cancer

Elsa Axelsdottir[1]

Long non-coding RNA (lncRNA) is an integral part of the transcriptome and provides a regulatory role in the inactivation of genes. By methylating chromatin, lncRNA can cause gene silencing and form heterochromatin. These transcripts can act both in-cis and in-trans; however, this review focuses on the specific lncRNA: HOX transcript antisense intergenic RNA (HOTAIR) in-trans. The long RNA sequence of HOTAIR binds to the HOX C locus and consequently silences chromatin at the HOX D locus. This silencing has shown to be indicative of specific cancers and their progression. In this review, HOTAIR's involvement in cancer is discussed. Here, evidence for the direct relationship between the over expression of HOTAIR and an individual's tumour progression into metastasis is presented. The abundance of the transcript in cancerous tissues can be used to diagnose the depth and progression of the tumours, as well as the general prognosis of the patient. HOX transcript antisense intergenic RNA's mechanism of action presents itself as a potential target for cancer therapeutics. Blocking the interaction between the associated proteins and the transcript holds potential for the treatment of tumours, and the inhibition of their progression into metastasis. Previously unexamined, this study focuses on HOTAIR over expression and compares the relationship in different cancers. Here, it is suggested that HOTAIR might serve as a general biomarker for neoplastic diseases. Moreover, further investigation of HOTAIR's presence and mechanism of action could result in less invasive therapies and diagnostic tests with direct applications in a clinical setting.

INTRODUCTION

The human genome project discovered that only about 20,000 genes encoded protein, while the remaining 99% of the genome was left non-coding and hence non-functional (Lander *et al*., 2001). The current opinion in genetics, however, suggests the opposite.

The functions of non-coding RNA have begun to emerge with compelling evidence of a role in gene regulation. The transcriptome is defined as the total RNA transcribed from DNA and includes both coding and non-coding transcripts. Within the non-coding transcript there are two sub divisions: long ncRNA (lncRNA) and short ncRNA (sncRNA), both of which are characterized physically by the sequence length of the transcript. The functional definition of the non-coding transcript is accepted as the genetic information that does not code for protein, the latter being the functional unit of life. Consequently this definition yields little importance to the non-coding portion of the genome. However, it cannot be overlooked that the non-coding transcript comprises 99% of the genomic information. Thus, it is no surprise that these non-coding transcripts do indeed possess functions that are important. The importance of these non-coding transcripts has only recently been elucidated (Dermitzakis *et al*., 2005).

The first lncRNA to be identified was H19 in 1990. Sedimentation analysis showed that the RNA was not associated with any translational machinery, suggesting it was not involved in coding for protein (Brannan *et al*., 1990). Soon after the discovery of H19, more lncRNAs were quickly identified including Xist, the X chromosome silencer, (Borsani *et al*., 1991; Brockdorff *et al*., 1991; Brown *et al*., 1991) and HOTAIR, a silencer of the HOX D locus (Rinn *et al*., 2007). Xist was discovered by slot blot analysis and northern blot, initially used to examine the presence of the Xist cDNA sequence present in inactive X chromosomes (Brown *et al*., 1991). John Rinn at Yale University discovered the unique transcript of RNA, HOTAIR, first by observing them in fibroblasts. Tiling microarray analysis of the HOX gene loci found that HOTAIR associates with HOX C and silences the chromatin on HOX D (Rinn *et al*., 2007).

When lncRNA was first identified, technologies were limited. However, the technologies for lncRNA identification have since dramatically improved. There are now several ways to identify non-coding transcripts. These include Crosslinking immunoprecipitation (CLIP) and High throughput sequencing of RNA Isolated by crosslinking immunoprecipitation (HITS-CLIP), deep RNA sequencing, tiling microarray analysis, and genomic selex (Li *et al*., 2012). These have all been modified and developed for the demands of the current era of next generation sequencing. The sequences of these ncRNA are available in databases such as Rfam, RNAdb and lncrnadb, the latter being a database that specifically annotates lncRNA.

Gene regulation can be considered the control of active and inactive expressions of the gene. This review considers the inactivation of genes caused by methylation and the alteration of chromatin states from chromatin to heterochromatin. Both small and long non-coding RNA are involved in gene regulation, either by activating or inactivating genes (Mattick & Makunin, 2006). Some form and associate with protein complexes that specifically methylate chromatin, while others act individually, silencing the gene by specific binding to the sequence. In addition, it is becoming evident that small and lncRNAs regulatory roles may be interdependent (Nana-Sinkam & Croce, 2011). LncRNA, in particular, has demonstrated a gene regulatory role in cancer, controlling the on and off states of tumour suppressor genes.

[1]University of Manchester, Oxford Road, Manchester M13 9PL, UK

*To whom correspondence should be addressed: elsa.axelsdottir@gmail.com

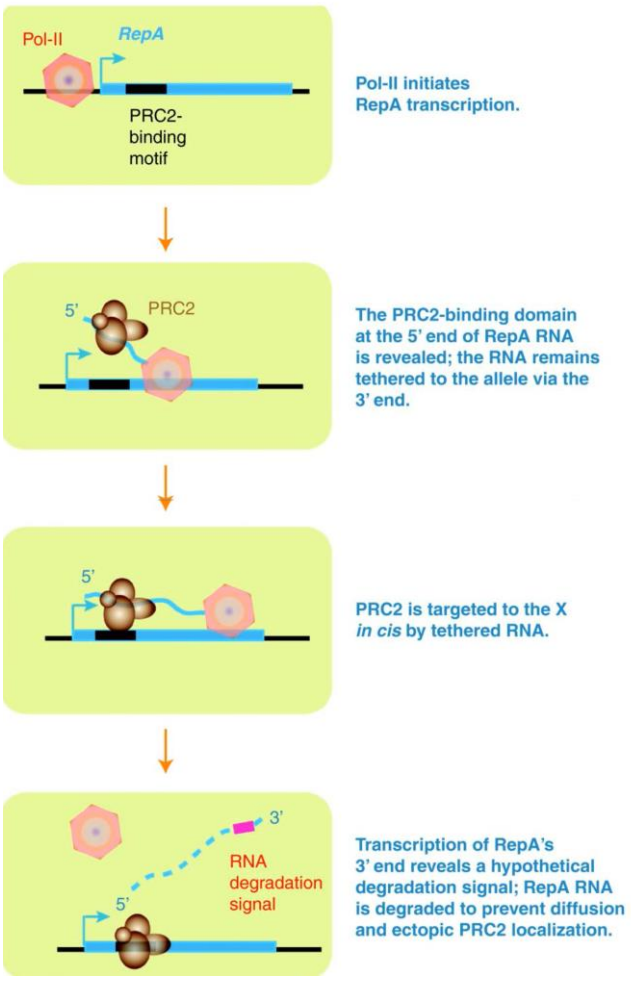

Figure 1. Mechanism of Xist RNA X chromosome inactivation by recruitment of PRC2. RepA RNA stays bound, allowing PRC2 recruitment to the X-chromosome. (Lee, 2010)

Long non-coding RNAs play a key role in gene regulation and pathology. In particular, HOTAIR plays a role in cancer. This review will investigate the role of HOTAIR in four types of cancer: breast, lung, liver, and colon. In addition, the possibility of HOTAIR as a diagnostic tool for cancer progression will be discussed. Furthermore, a mechanistic investigation of HOTAIR could allow for the discovery of future therapeutic cancer targets.

Long non-coding RNA (lncRNA)

Long non-coding RNAs are most often defined as sequences starting from 200 bases to about 100kb (Setoyama *et al*., 2011). These long transcripts can take many different forms and include long intergenic non-coding RNA (lincRNA) as well as transcribed ultra-conserved regions (T-UCR) (Esteller, 2011). Examples of these and their functions are shown in Table 1. It is important to note that other types of lncRNAs in humans are also present, including ribosomal RNAs and pseudogenes (Gibb *et al*., 2011)

Long non-coding transcripts are unique in that they can be classified in several different ways including by function, mechanism, loci of origin, and orientation (Ponting *et al*., 2009; Li *et al*., 2012). The first classification is by function, that is, whether they are regulatory or not (Ponting *et al*., 2009). This requires the distinction between regulatory function and transcriptional noise(Ponting *et al*., 2009). Therefore there must be a distinction between the functional and non-functional sequences. Classification by mechanism defines the different lncRNAs that alter structure by either discrete binding to the DNA sequence itself or by recruiting intermediates. When lncRNA binds directly to the DNA, the mechanism of action is referred to as in-cis. In this event the chromatin structure is directly modified by the presence of the bound lncRNA molecule. When the lncRNA instead recruits intermediate protein complexes to mediate changes in chromatin structure, the mechanism is referred to as in-trans. The process of modification in-trans is known as transvection (Mercer *et al*., 2009) and will be the focus of this review. Further classification by loci of origin and orientation defines the transcripts by the physical location within the gene i.e. intergenic, partially intronic and totally intronic (Nakaya *et al*., 2007) and location on either sense or antisense strands (He *et al*., 2008) respectively. The mechanism of chromatin alteration by lncRNA is governed by the lncRNA molecule's ability to stabilize protein complexes for the process of chromatin alteration and may also provide a scaffold for chromatin alteration of specific genes (Shamovsky & Nudler, 2006). The stability of lncRNA may also affect the stability of other RNA sequences (Ramaiah *et al*., 2012). It has been found that in fact antisense sequences may play a role in the regulation of sense sequences and seem to negatively regulate the sense transcripts (Lapidot & Pilpel, 2006; He *et al*., 2008).

Classifying the transcripts serves to organize newly discovered lncRNA and allows for the grouping of RNAs into families. The respective functions can then be inferred accordingly. Rfam is a database that annotates these sequences (Griffiths-Jones *et al*., 2003).

Name	Size	Location	Number in Humans	Functions	Examples
LincRNA	>200 bp	Widespread loci	>1,000	DNA-chromatin complex formation	HOTAIR, HOTTIP, lincRNA-p21
T-UCR	>200bp	Widespread loci	>350	Possibly the regulation of mRNA and miRNA	Uc.283+, uc.338, uc160+
Other	>200bp	Widespread loci	>3,000	Imprinting, X chromosome inactivation, telomere regulation	XIST, TSIX. TERRAs, p15AS. H19. HYMAI

Table 1. Examples of lncRNAs in humans. (Adapted from Esteller, 2011)

Mechanism of chromatin alteration in-cis

There are several well-identified and characterized examples of chromatin alterations in-cis, despite a majority that remain to be discovered. Chromatin modification by non-coding RNA in-cis can be described through the examples of two long non-coding RNAs: Xist and AIR.

Xist is a lncRNA transcript involved in X-chromosome inactivation. The X-chromosome inactivation centre (Xic) locus, characterized by a repeated A motif encoding an RNA called RepA, is involved in the initiation of X-chromosome inactivation. Regulatory lncRNAs are present at this centre and mediate the inactivation in-cis (Rastan, 1994). The polycomb repressor group protein, PRC2, is suggested to associate with Xist to help the process of X-chromosome inactivation (Li *et al*., 2012). With an accumulation of Xist lncRNA, PRC2 proteins are recruited to the X-chromosome with the help of RepA (Figure 1) (Lee, 2010). The binding of Xist in-cis to cause methylation of the chromatin by PRC2 silences transcription and inactivates the X-chromosome (Li *et al*., 2012).

The lncRNA, Antisense Igf2r RNA (AIR), epigenetically silences the slc22a3 gene in-cis. Antisense Igf2r RNA is an antisense transcript that is associated with chromatin modifying complexes PRC2 and G9a (Nagano *et al*., 2008). G9a is a chromatin modifying complex that trimethylates lysine 9 (Khalil *et al*., 2009). AIR is transcribed from the hypomethylated paternal allele on the Insulin like growth factor 2 receptor (IGF2R) gene (Della Vedova & Cone, 2004). AIR directly overlaps the slc22a3 gene in the antisense orientation to induce silencing (Mohammad *et al*., 2009).

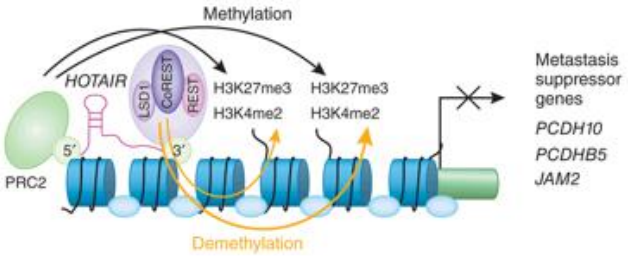

Figure 2. HOTAIR forms a bridge between polycomb repressive complex (PRC2) and histone lysine demethylase (LSD1), creating a molecular scaffold for the protein complexes to induce post-translational modifications (PTMS). PRC2 mediates PTMs while LSD1 inhibits them. (De Lucia & Dean, 2011)

Mechanism of chromatin alteration in-trans

HOX transcript antisense intergenic RNA (HOTAIR) is one of the few well-characterized examples of lncRNA action in-trans. HOX transcript antisense intergenic RNA is an antisense long intergenic non-coding RNA (lincRNA). This lncRNA is classified as a HOX locus RNA, meaning its location of transcription as well as the location where it silences chromatin are both within the HOX loci (Rinn *et al*., 2007). Its length is 2.2kb and is transcribed from the HOXC locus, located on human chromosome 12, intergenic of HOXC11 and HOXC12. HOTAIR then silences the chromatin of HOX D, 40kbps away on chromosome 2 (Rinn *et al*., 2007). The activity of the lincRNA between two HOX loci illustrates the mechanism referred to as "in-trans".

HOX loci are a series of transcription factors that are important for segmental identity and body patterning (Myers, 2008). HOX genes determine positional identity (Chang, 2009). They are often referred to as 'master control genes' because of their large developmental importance in identifying the axes in development. It is hypothesized that all HOX genes are regulated through either a cis or trans mediated lncRNA (Prensner & Chinnaiyan, 2011). There is evidence suggesting that the majority of lncRNA in the HOX loci regulates the genes in-trans (Khalil *et al*., 2009).

HOX loci chromatin can be present in either ON or OFF states, depending upon the concerned differentiated cells. These different states are programmed by two groups of proteins known as trithorax group proteins and polycomb group proteins (PcGs) (Chang, 2009). Trithorax group proteins program the ON state of chromatin and induce tri-methylation of lysine 4 on histone 3, allowing histone 4 to be acetylated. Polycomb group proteins, however, are antagonistic to the trithorax programming and instead induce the inactive chromatin state by trimethylation of lysine 27 on histone 3, inhibiting acetylation of histone 4 and hence maintaining the silent state of chromatin (Schuettengruber *et al*., 2007). The HOX loci may transcribe several lncRNAs, highlighting them as global regulators (Khalil *et al*., 2009). Many human lincRNAs associate with chromatin-modifying complexes and affect gene expression.

HOTAIR does not associate with DNA alone, but also forms a bridge between protein complexes, forming a molecular scaffold to support chromatin silencing. These complexes of proteins include the Polycomb repressor complex 2 (PRC2), histone lysine demethylase (LSD1), co-repressor for elements-1-silencing transcription factor (CoREST), and repressor for elements-1-silencing transcription factor (REST). LSD1, CoREST, and REST are a complex of three interacting proteins (Tsai *et al*., 2010). HOTAIR is associated with PRC2 at the 5' end and LSD1, CoREST and REST at the 3' end (Figures 3 and 4) (Tsai *et al*., 2010). LSD1 is a protein that forms a complex with CoREST that then bridges to REST, a neuronal gene silencer, and mediates silencing of the target gene (Figure 3) (Tsai *et al*., 2010). PRC 2 is made up of protein subunits EZH2, SUZ12, and EED (Kogo *et al*., 2011). EZH2 is an H3k27 methylase and is phosphorylated by the cyclin dependent kinase CDK1, which then allows HOTAIR to bind to the protein and the rest of the PRC2 complex associates, allowing association with the histone. HOTAIR binds to EZH2 and then recruits and associates with the PRC2 complex (Kaneko *et al*., 2010). Polycomb grouping proteins are recruited to perform the chromatin alteration (Simon & Kingston, 2009), with most lncRNA associating with PRC2 (Khalil *et al*., 2009).

It is still unclear whether HOTAIR binds for the purpose of locating the chromatin modifying complexes or if they are also involved in the activation of the protein complexes, exhibiting an effector's mechanism (Wang *et al*., 2011).

However, it is clear that HOTAIR binds to these protein complexes prior to the binding onto DNA. It is the coordinated protein complex that binds to and associates with the DNA which leads to inductions of methylation and silences the gene, forming heterochromatin (Tsai *et al*., 2010). Hence, HOTAIR is used as a foundation for polycomb remodelling complexes and histone modification (Prensner & Chinnaiyan, 2011).

HOTAIR helps to induce a specific pattern of methylation, which involves trimethylation as well as demethylation (Croce, 2010). Trimethylation of Histone 3 lysine 27 is mediated by PRC2, while Histone 3 lysine 4 demethylation is mediated by the LSD1 complex (Figure 3) (Croce, 2010). Although the pattern of methylation mediated by HOTAIR is as stated above, the transcription of HOTAIR can also be identified by another specific methylation pattern. This methylation pattern is K4-K36 tri-methylation sites on histone 3 and is an important marker for the presence of the transcription of HOTAIR. This methylation pattern indicates that the transcription of HOTAIR is active and is thus present in the cell. Therefore, it is inferred that a K4-K36 methylated domain is a chromatin signature that indicates the presence of HOTAIR (Khalil *et al*., 2009).

HOTAIR in Cancer

HOTAIR is a lincRNA that has been found in several forms of cancer. The mechanism of HOTAIR remains consistent across different cancer types, mainly involving altering of the HOX D locus. Here four cancers will be discussed; breast cancer, lung cancer, hepatocellular cancer and colorectal cancer. Although these four present the most compelling evidence for the presence of HOTAIR, it is unlikely that the cancers discussed here construct a comprehensive list. This case study provides evidence for HOTAIR's ability to promote the localization of the polycomb repressive complex and trimethylation of H3K27 as well as progression into metastasis (Khalil *et al*., 2009). Evidence for the expression of HOTAIR in tumorigenesis and metastasis is discussed.

It is important to note that HOTAIR has low sequence conservation between species, but conservation is nonetheless still present (Rinn *et al*., 2007; Prensner & Chinnaiyan, 2011). HOTAIR only exists in mammals and may differ between species (Gupta *et al*., 2010). It is hypothesised that HOTAIR has gained a functional importance in humans due to a rapid evolution. HOTAIR is a term that is used to refer to the human transcript, while the mouse HOTAIR is denoted as mHOTAIR (Schorderet & Duboule, 2011). Here, the human HOTAIR sequence was assessed in mouse (Gupta *et al*., 2010).

Breast Cancer

Breast cancer is one of the most common cancers among women and involves the growth of tumours in the breast epithelial tissue (Jemal *et al*., 2011). Several studies support the role of HOTAIR in the cause and progression of breast cancer (Prensner & Chinnaiyan, 2011). In normal tissues, many non-coding RNAs act as tumour suppressors and protect the cells from cancer development (Croce, 2010). The protective mechanism is through the use of tumour suppressor genes. Some HOX lincRNAs are breast tumour suppressors, such as lncRNA HOX5 (Gupta *et al*., 2010). The expression of these lincRNAs is deregulated in cancerous tissue (Raman *et al*., 2000). HOTAIR specifically silences a variety of these tumour suppressor genes namely the protocadherin family such as HOXD10 and PGR, as well as several metastasis suppressor genes including PCDH10, PCDHB5, and JAM2 (Figure 4) (Croce, 2010; Gupta *et al*., 2010; Kogo *et al*., 2011). Hence, the cells are not protected against uncontrolled division and cancer can ensue. HOTAIR is upregulated in cancer and this misregulation of the lincRNA has been proven to be detrimental to the prognosis of patients (Gupta *et al*., 2010). The transformation of breast epithelial cells is also associated with the EZH2 subunit, which is an enhancer of zeste homolog in humans. Its presence is similarly associated with the progression of breast cancer (Kleer *et al*., 2003).

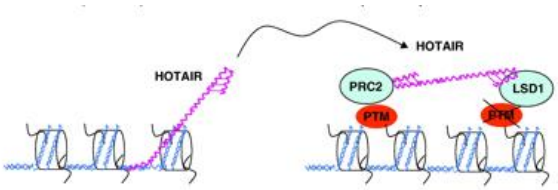

Figure 3. Mechanism of breast cancer metastasis by HOTAIR. The PRC2 complex contains the subunits EZH2, SUZ12, EED associate with lincRNA HOTAIR and form a bridge with LSD1, CoREST, and REST inducing methylation and heterochromatin formation. This inhibits the transcription of the metastasis suppressor genes, PCDH10, PCDHB5, and JAM2, allowing metastasis to ensue. (Croce, 2010)

Additionally, the expression of HOTAIR in cancerous tissues increases the mobility of the cells. The nature of cancer allows the tumorous cells to move and invade other tissues, causing a phenomenon known as metastasis. There is a link between metastasis and over expression of HOTAIR as it is suspected to silence multiple metastasis suppressor genes (Gupta *et al*., 2010; Kogo *et al*., 2011). HOTAIR associates with and induces tri-methylation of these metastasis suppressor genes (Gupta *et al*., 2010). HOTAIR induces H3K27me3 and increases PRC2 expression in breast cancer cells (Gupta *et al*., 2010; Kogo *et al*., 2011). It was observed that these cells with the H3K27me3 methylation profile resemble embryonic fibroblast cells (Gupta *et al*., 2010; Gibb *et al*., 2011). The nature of this reprogramming causes the metastatic mobile and invasive properties. HOTAIR expression is linked to a poor prognosis of breast cancer patients (Gupta *et al*., 2010). Although a lot of work has been done on HOTAIR's presence in breast cancer, it is becoming increasingly evident that this lincRNA plays a regulatory role in a number of other cancers such as lung, hepatocellular, and colon cancer as well.

Lung Cancer

Many cancers metastasize to the lung, but it has been found that cancerous tissues with the over expression of HOTAIR tend to

increase their mobility to the lung. In particular, the movement of cancerous cells from breast to the lung is discussed in Gupta's paper (Gupta *et al*., 2010). The non-coding RNA MALAT-1 had previously characterized lung cancer, but evidence for the presence of HOTAIR in lung metastasis from the breast is strong, raising the possibility of it playing an important role as well (Gibb *et al*., 2011). When a non-metastatic cell line was induced with the expression of HOTAIR, the cell line was able to move to and invade the lung (Gupta *et al*., 2010; Liu *et al*., 2013).

Hepatocellular Cancer

Hepatocellular carcinoma (HCC) presents itself with over 600,000 cases per year worldwide. One growing problem with HCC is that patients treated with a liver transplant often suffer from tumour relapse and ultimately progression into metastasis. Thus the understanding of metastasis and cancer progression is important. In hepatocellular cancer, again, the over expression of HOTAIR is linked to cancerous tissue. Interestingly, the presence of HOTAIR in the tissue is linked to recurrence and relapse in patients (Yang *et al*., 2011). This is indicative of the transcripts ability to migrate and invade. HOTAIR expression was lower in non-cancerous tissue than in cancerous tissue (Geng *et al*., 2011; Yang *et al*., 2011). The cumulative relapse-free survival of patients with high HOTAIR expression was significantly lower than in patients with low expression of the lincRNA, as analysed by the Kaplan-Meier method and log-rank test (Geng *et al*., 2011). Additionally, the cancerous tissue was more responsive to chemotherapy and became more sensitive to apoptosis mediated by TNF-a with lower expression of HOTAIR (Yang *et al*., 2011). The EZH2 subunit of PRC2 is also important in HCC, with high levels of expression in HCC of human cell lines (Sudo *et al*., 2005). The confirmation of EZH2's importance in the progression of cancer indicates the importance of the collective interactions between the subunits that allow cancerous properties to progress.

Colorectal Cancer

In a single study on colorectal cancer (CRC), it was found that the state of the tumour was related to the expression of HOTAIR. The depth of the tumour and differentiation of the cells were also related to the expression of HOTAIR. High HOTAIR expression was associated with metastasis in the liver. This indicates that the expression of HOTAIR and invasiveness of the cancerous cells are directly related. Increasing HOTAIR expression increased invasion, while decreasing HOTAIR expression decreased invasive properties. Therefore, it is likely that HOTAIR may be responsible for increasing the number of undifferentiated cancer cells. The increased expression of HOTAIR was an independently causative variable as invasive properties increased (Kogo *et al*., 2011). Further study on the invasive properties of colorectal cancer cells will be required to determine the role of HOTAIR as a causative factor in the progression of the disease.

Diagnosis and Therapy

Cancer is a misregulation of pathways within the cell and lncRNAs are often associated with this misregulation in the cancerous tissue. There is a clear trend that the levels of expression are linked to cancer progression. Hence there are emerging proposals of the use of lncRNAs as biomarkers for cancer (Tsai *et al*., 2011). As discussed above, there is a strong relationship between the levels of lncRNAs in the cell and subsequent metastasis. This may have profound clinical implications. As seen in hepatocellular cancer, HOTAIR is not only a biomarker for the progression of cancer but also a marker for the potential relapse, indicating the transcript could be important in cancer risk assessment as well as in the process of classifying patient prognosis. The use of HOTAIR as a biomarker in hepatocellular cancer could prove to be clinically significant, particularly in patients undergoing liver transplants (Yang *et al*., 2011). Increased levels of HOTAIR can help estimate relapse-free survival in liver cancer patients. Similarly, it would be just as significant to use this relationship to determine the potential for metastasis from breast to lung as well as depth of the tumour in colorectal cancer (Gupta *et al*., 2010; Kogo *et al*., 2011). It seems clear that the use of HOTAIR as a biomarker can have immediate benefits if incorporated into clinical practice as a diagnostic tool. The interaction between HOTAIR and the two protein complexes, PRC2 and LSD1, are also instrumental in methylating the chromatin and as a consequence, they could also be used for diagnosing the methylation that occurs. A direct relationship between the HOTAIR associated protein expression and methylation is apparent: a high expression of PRC2 and LSD1 yields a high level of chromatin methylation and gene silencing by HOTAIR (Sudo *et al*., 2005). This relationship can be used as a diagnostic tool to further clarify the diagnosis.

In addition, breast cancer in particular is known to be highly resistant to chemotherapy and radiotherapy (Mego *et al*., 2010). This is significant in that there is a need for alternative methods to treat the cancer. It is well characterized that the PRC2 complex plays a fundamental role in the mechanism of HOTAIR mediated chromatin silencing and thus this could be a target for therapy. Exploiting the interaction of PRC2 and HOTAIR may interfere with the functional process of the chromatin silencing. Small molecule interference as well as oligonucleotide interference have been methods of interfering with small ncRNA and could therefore have the potential to work with the interference of HOTAIR and its respective protein subunits. A possible method of this interference could be the introduction of a competitive antagonist. Additionally, fragmenting large ribonucleoprotein complexes would inhibit the formation of the HOTAIR and EZH2 complex (Tsai *et al*., 2011) and is, thus, a potential therapeutic strategy. In addition, developing a method of reversing the methylation mediated by HOTAIR by introducing demethylase complexes or introducing trithorax group proteins to counteract the effects of the PRC2 protein may re-activate the tumour suppressor genes.

It is significant to note that it is not only the mechanism but also the misregulation of lincRNAs that may be important to cancer progression (Gupta *et al*., 2010). Therefore, maintaining the function of the regulatory non-coding RNA would be an important task for therapeutics. This could be achieved by

suppressing the transcription of HOTAIR potentially by RNA interference (Tsai *et al*., 2011).

CONCLUSION

These four examples present compelling evidence for the importance of HOTAIR in the phenomenon of cancer and it is for this reason that the expression pattern of long non-coding RNA is rapidly becoming accepted as a hallmark of the disease (Gibb *et al*., 2011). Screening for the over expression of this transcript in patients' cells can be used to identify cancer progression. Hence, HOTAIR can be used as a biomarker for cancer diagnosis. The current mechanistic understanding gives rise to potential therapies to target the functions of HOTAIR. Applications of such therapies are yet to be implemented, but a movement away from chemotherapy and radiation towards the direct targeting of HOTAIR could potentially have profound impacts on the treatment of cancer and patient survival. .

HOX transcript antisense intergenic long non-coding RNA, however, is not the only lincRNA transcribed at the HOX locus. There are still many aspects to the regulatory mechanism of ncRNA yet to be identified. Most lincRNAs use the PRC2 pathway, hence, HOTAIR can be used as a model to help understand the mechanism of other lncRNAs (Khalil *et al*., 2009).

It would be interesting to identify the metastatic trends between tissue types affected by tumours with cancer. As there are several other non-coding transcripts that are prevalent in cancer, examining the associations and expression profiles of each transcript in cancer would greatly increase the understanding of the importance of lncRNA in the disease. In particular MALAT-1 is known to associate with lung cancer (Gibb *et al*., 2011) and it would be interesting to see if MALAT- 1 was also present in the metastatic lung tissue in Gupta's breast cancer patients. This could help elucidate the interaction of ncRNAs together, as cancer is a multi-factorial disease.

Additionally, it is possible that the functions of both small and lncRNAs are intertwined (Nana-Sinkam & Croce, 2011). Understanding the interdependence between the transcripts will be important for gaining a comprehensive understanding of the mechanism in which the genes are silenced as well as the potential treatment targets. It is important to the progression of the field that a complete list of ncRNA for bio-marking in cancer be identified (Huarte & Rinn, 2010).

HOX transcript antisense intergenic RNA is becoming increasingly well-known and abundant in several tissue types and it is unlikely to be limited to the tissue and cancer types discussed in this review, therefore further exploration is required to enhance the understanding of the mechanistic action of lncRNA in cancer.

REFERENCES

Baker M. (2011). Long noncoding RNAs: the search for function. *Nature Methods* **8,** 379-383.

Borsani G, Tonlorenzi R, Simmler MC, Dandolo L, Arnaud D, Capra V, Grompe M, Ballabio A. (1991). Characterization of a murine gene expressed from the inactive X-chromosome. *Nature* **351,** 325-329.

Brannan CI, Dees EC, Ingram RS & Tilghman SM. (1990). The product of the H19 gene may function as an RNA. *Molecular and Cellular Biology* **10,** 28-36.

Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, McCabe VM, Norris DP, Penny GD, Patel D & Rastan S. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X-chromosome. *Nature* **351,** 329-331.

Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R & Willard HF. (1991). A gene from the region of the human X-inactivation center is expressed exclusively from the inactive X-chromosome. *Nature* **349,** 38-44.

Chang HY. (2009). Anatomic Demarcation of Cells: Genes to Patterns. *Science* **326,** 1206-1207.

Croce CM. (2010). LINCing chromatin remodeling to metastasis. *Nature Biotechnology* **28,** 931-933.

Della Vedova CB & Cone KC. (2004). Paramutation: The chromatin connection. *Plant Cell* **16,** 1358-1364.

Dermitzakis ET, Reymond A & Antonarakis SE. (2005). Conserved non-genic sequences - an unexpected feature of mammalian genomes. *Nature Reviews Genetics* **6,** 151-157.

Esteller M. (2011). Non-coding RNAs in human disease. *Nature Reviews Genetics* **12,** 861-874.

Geng YJ, Xie SL, Li Q, Ma J & Wang GY. (2011). Large Intervening Non-coding RNA HOTAIR is Associated with Hepatocellular Carcinoma Progression. *The Journal of international medical research* **39,** 2119-2128.

Gibb EA, Brown CJ & Lam WL. (2011). The functional role of long non-coding RNA in human carcinomas. *Molecular Cancer* **10**.

Griffiths-Jones S, Bateman A, Marshall M, Khanna A & Eddy SR. (2003). Rfam: an RNA family database. *Nucleic Acids Research* **31,** 439-441.

Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, … Chang HY. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464,** 1071-U1148.

He YP, Vogelstein B, Velculescu VE, Papadopoulos N & Kinzler KW. (2008). The Antisense Transcriptomes of Human Cells. *Science* **322,** 1855-1857.

Huarte M & Rinn JL. (2010). Large non-coding RNAs: missing links in cancer? *Human Molecular Genetics* **19,** R152-R161.

Jemal A, Bray F, Center MM, Ferlay J, Ward E & Forman D. (2011). Global Cancer Statistics. *Ca-a Cancer Journal for Clinicians* **61**.

Kaneko S, Li G, Son J, Xu CF, Margueron R, Neubert TA & Reinberg D. (2010). Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes & Development* **24,** 2615-2620.

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Morales DR, Thomas K, … Rinn JL. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 11667-11672.

Kleer CG, Cao Q, Varambally S, Shen RL, Ota L, Tomlins SA, Ghosh D, … Chinnaiyan AM. (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 11606-11611.

Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, … Mori M. (2011). Long Noncoding RNA HOTAIR Regulates Polycomb-Dependent Chromatin Modification and Is Associated with Poor Prognosis in Colorectal Cancers. *Cancer Research* **71,** 6320-6326.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, … Int Human Genome Sequencing C. (2001). Initial sequencing and analysis of the human genome. *Nature* **409,** 860-921.

Lapidot M & Pilpel Y. (2006). Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *Embo Reports* **7,** 1216-1222.

Li X, Wu Z, Fu X & Han W. (2012). Long Noncoding RNAs: Insights from Biological Features and Functions to Diseases. *Medicinal Research Reviews*.

Liu, X., Liu, Z., Sun, M., Liu, J., Wang, Z., De, W. (2013). The long non-coding RNA indicates poor prognosis and promotes metastasis in non-small cell lung cancer. *BMC Cancer* **13**:464.

Mattick JS & Makunin IV. (2006). Non-coding RNA. *Human Molecular Genetics* **15,** R17-R29.

Mego M, Mani SA & Cristofanilli M. (2010). Molecular mechanisms of metastasis in breast cancer-clinical applications. *Nature Reviews Clinical Oncology* **7,** 693-701.

Mercer TR, Dinger ME & Mattick JS. (2009). Long non-coding RNAs: insights into functions. *Nature Reviews Genetics* **10,** 155-159.

Mohammad F, Mondal T & Kanduri C. (2009). Epigenetics of imprinted long noncoding RNAs. *Epigenetics* **4,** 277-286.

Myers P. (2008). Hox genes in development: The Hox code. Nature Education.

Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R & Fraser P. (2008). The Air Noncoding RNA Epigenetically Silences Transcription by Targeting G9a to Chromatin. *Science* **322,** 1717-1720.

Nakaya HI, Amaral PP, Louro R, Lopes A, Fachel AA, Moreira YB, El-Jundi TA, … Verjovski-Almeida S. (2007). Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. *Genome Biology* **8**.

Nana-Sinkam SP & Croce CM. (2011). Non-coding RNAs in cancer initiation and progression and as novel biomarkers. *Molecular Oncology* **5,** 483-491.

Negrini M, Nicoloso MS & Calin GA. (2009). MicroRNAs and cancer-new paradigms in molecular oncology. *Current Opinion in Cell Biology* **21,** 470-479.

Ponting CP, Oliver PL & Reik W. (2009). Evolution and Functions of Long Noncoding RNAs. *Cell* **136,** 629-641.

Prensner JR & Chinnaiyan AM. (2011). The Emergence of lncRNAs in Cancer Biology. *Cancer Discovery* **1,** 391-407.

Ramaiah M, Shum EY & Wilkinson MF. (2012). How to Activate a Gene: Decap Its Associated Noncoding RNA. *Molecular Cell* **45,** 271-273.

Raman V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E, Marks J & Sukumar S. (2000). Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* **405,** 974-978.

Rastan S. (1994). X chromosome inactivation and the Xist gene. *Current Opinion in Genetics and Development* **4,** 292-297.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, … Chang HY. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by Noncoding RNAs. *Cell* **129,** 1311-1323.

Schorderet P & Duboule D. (2011). Structural and Functional Differences in the Long Non-Coding RNA Hotair in Mouse and Human. *Plos Genetics* **7**.

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B & Cavalli G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* **128,** 735-745.

Setoyama T, Ling H, Natsugoe S & Calin g. (2011). Non-coding RNAs for Medical Practice in Oncology. *Keio Journal of Medicine* **60,** 106-113.

Shamovsky I & Nudler E. (2006). Gene Control by Large Noncoding RNAs. *Science* **2006**.

Simon JA & Kingston RE. (2009). Mechanisms of Polycomb gene silencing: knowns and unknowns. *Nature Reviews Molecular Cell Biology* **10,** 697-708.

Sudo T, Utsunomiya T, Mimori K, Nagahara H, Ogawa K, Inoue H, Wakiyama S, … Mori M. (2005). Clinicopathological significance of EZH2 mRNA expression in patients with hepatocellular carcinoma. *British Journal of Cancer* **92,** 1754-1758.

Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E & Chang HY. (2010). Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science* **329,** 689-693.

Tsai MC, Spitale RC & Chang HY. (2011). Long Intergenic Noncoding RNAs: New Links in Cancer Progression. *Cancer Research* **71,** 3-7.

Wang X, Song X, Glass CK & Rosenfeld MG. (2011). The Long Arm of Long Noncoding RNAs: Roles as Sensors Regulating Gene Transcriptional Programs. *Cold Spring Harbor Perspectives in Biology* **3**.

Yang Z, Zhou L, Wu LM, Lai MC, Xie HY, Zhang F & Zheng SS. (2011). Over expression of Long Non-coding RNA HOTAIR Predicts Tumor Recurrence in Hepatocellular Carcinoma Patients Following Liver Transplantation. *Annals of Surgical Oncology* **18,** 1243-1250.