Abstract. Recently, the homeobox (HOX) gene family has been reported as a factor in tumorigenesis. In the human genome, the HOX gene family contains 4 clusters with 39 genes and multiple transcripts. Mutation or abnormal expression of genes is responsible for developmental disorders. In addition, changes in the levels and activation of certain HOX genes has been associated with the development of cancer. Long non-coding RNAs (lncRNAs) have also been identified to serve critical functions in cancer. Although a limited number of lncRNAs have been previously investigated, the list of functional lncRNA genes has recently grown. Two of the most important and well-studied lncRNAs and HOX transcript genes are HOX transcript antisense RNA (HOTAIR) and HOXA distal transcript antisense RNA (HOTTIP). The present study aimed to review not only the function of the HOTAIR and HOTTIP genes in certain forms of cancer, but also to review other HOX genes and protein functions in cancer, particularly HOX family genes associated with lncRNAs.

1. Introduction

The homeobox (HOX) gene family is a large group of comparable genes that function during early embryonic development to direct the formation of various body structures (1,2). In humans, it is estimated that the HOX gene family contains 65 pseudogenes and 235 functional genes (3). HOX genes are present in every human chromosome, and are commonly organized in clusters. A number of subfamilies and classes of HOX genes have been described, although these categories are used inconsistently.

Various genes of the HOX family are involved in a broad range of crucial developmental activities (4-7). Tumor suppressing regions of HOX genes have been demonstrated to inhibit cell growth in an uncontrolled manner (4). These genes commonly perform important functions, and mutation or abnormal expression of these genes has been associated with developmental disorders and several forms of cancer (5-7).

In the human genome, only 2% of the sequence encodes proteins that function in biological progression (8), with ~90% consisting of non-coding RNAs (ncRNAs), which are described as the ‘noise’ of the genome (9-12). Over the last few years, an increasing number of long ncRNAs (lncRNAs) have been investigated by microarray and sequencing technology, and it has been proposed that they may regulate various major biological processes involved in metabolism, differentiation and tumor development (13). ncRNAs are divided into two subgroups, including small ncRNAs (snRNAs) and lncRNAs. ncRNAs are defined as snRNAs when the RNA length is <200 nt, which includes microRNAs (also named as miRNAs/miRs). Conversely, lncRNAs contain >200 nt of transcript. Currently, lncRNAs are becoming more popular and attracting more attention in scientific research (14-16).

Two important platforms have been used for analysis of lncRNAs; advanced sequencing platforms estimated that there are ~20,000 distinct IncRNAs in humans (17,18). The other platform is the commercial lncRNAs ChIP, consisting of ~30,000 IncRNA transcripts, and is an ideal tool for making comparisons between tumorous and non-tumorous samples (19,20). It has been suggested that IncRNAs have a critical role in cancer development. However, the biological functions of a large proportion of IncRNAs remain to be elucidated. Previous studies reported that IncRNAs are involved in fundamental cellular processes, including apoptosis, cell cycle and cell proliferation (21-23), chromatin modification (24,25),
genomic reprogramming (26,27), gene imprinting (28) and RNA processing (29). It was demonstrated that tumor or non-tumor tissues exhibited different profiles in various forms of human cancer, and altered lncRNA expression has been functionally linked to tumorigenesis; thus suggesting that certain lncRNAs function as tumor promoters or suppressors (17,30,31). Various studies have determined that certain lncRNAs may be detected through their altered expression in different forms of cancer (32-35). The identification of cancer-associated lncRNAs (CA-lncRNAs), and interactions between the CA-lncRNAs and target genes, are beneficial to supplement the current knowledge of signaling pathways and are conducive to a better understanding of cancer development.

One of the most important and well-studied lncRNA and HOX transcript gene is HOTAIR, which is located on chromosome 12 within the HOXC gene cluster and is 2,158 nt long (32,36). HOTAIR is unique in that it is overexpressed in the vast majority of cancer types and has been recognized as an oncogenic lncRNA (32). Recently, it was reported that HOTAIR induced proliferation and metastasis in a variety of tumors and functioned as a negative prognostic indicator for several forms of cancer (32,36). In addition, it was observed that the lncRNA interacted with polycomb repressive complex 2 (PRC2) and inhibited multiple genes, particularly metastasis-suppressing genes, in cancer tissues (36,37). Subsequent studies demonstrated that HOTAIR serves as a molecular scaffold for at least two distinct histone modification complexes, coordinating their functions in transcription repression (25). Several transcriptome-wide association studies have detected extensive alterations in cellular transcript levels in response to HOTAIR inhibition, indicating that HOTAIR may regulate hundreds of genes (33-35), thus providing insight into the underlying mechanisms of this lncRNA in cancer. Although transcript frequencies are informative, they do not necessarily reflect the level of cellular protein; this is due to the fact that protein activity is affected by an array of post-transcriptional regulatory mechanisms and the association between mRNA and protein levels is generally modest (34,38).

LncRNAs have been frequently investigated in recent years and certain lncRNAs associated with HOX genes have been detected in sequencing and ChIP platform research. It is expected that an increasing number of HOX gene-associated lncRNAs may be identified in the coming years, which will further clarify the association between lncRNAs and HOX in tumorigenesis. The present review discusses HOX genes and research regarding their association with lncRNAs, particularly in the context of carcinogenesis.

2. HOX genes and cancer

HOX genes were originally identified in Drosophila melanogaster (39). These genes most importantly function as transcriptional regulators during normal morphogenesis in the process of cell-to-cell communication, the modification of which may contribute to the development of cancer (40,41). The HOX gene homology domain is able to bind to specific DNA sequences and regulate gene transcription (42). However, possible mechanisms underlying the function of HOX genes in tumorigenesis have not yet been elucidated. Notably, abnormal expression of various HOX genes has been observed in a number of solid tumors and hematological malignancies (43-46), and studies have begun to distinguish the biological mechanisms associated with their expression (47-50). The human HOX gene family contains 4 HOX gene clusters, which are presented in Table I.

A study by Hur et al (51) profiled 39 HOX genes involved in breast cancer, but only 25 of these genes were examined in malignant and non-malignant tissues. A total of 14/25 genes exhibited differential expression levels in malignant breast tissues compared with non-malignant breast tissues (51). Several studies have demonstrated that various HOX genes are differentially expressed (either up- or downregulated) in patients with breast cancer, indicating the involvement of these genes in carcinogenesis and breast cancer metastasis (43,52,53). In addition, HOXA5 and HOXA9 have been implicated as regulators of p53 and breast cancer 1 (BRCA1), respectively, which are important factors in cancer development (54,55). Furthermore, the function of HOXA1, -B7 and -B9 in breast tumorigenesis has been investigated; HOXB7 was identified as a key factor regulating cell proliferation, invasion and tamoxifen resistance (56-58), and the overexpression of HOXA1 and HOXB9 contributed to breast cancer tumorigenicity (59,60). Additional studies have reported an association between poor prognosis and overexpression of HOXD3 and HOXB13 (61,62).

Accumulating evidence indicates that various tumors exhibit low levels of HOXA11 expression, which has been described as a tumor suppressor (65). A study by Bai et al (64) observed that the methylation frequencies of HOXA11 in gastric cancer tissues and adjacent cancer tissues were higher than those in normal gastric mucosa (P<0.05). Furthermore, HOXA11 expression was downregulated following hypermethylation of the promoter region (64); this suggests that HOXA11 may function as an important repressor in gastric cancer tumorigenesis, and aberrant promoter methylation may be the primary cause behind the loss or downregulation of HOXA11 expression, subsequently resulting in gastric cancer carcinogenesis. The study also observed that hypermethylation of HOXA11 was significantly associated with lymph node metastasis and Tumor-Node-Metastasis stage in gastric cancer, and in vitro experiments demonstrated that HOXA11 may control cell growth with its defect enhancing cell proliferation (64).

HOXC11 expression has been observed to be significantly higher in renal cell carcinoma (RCC) compared with normal kidney tissues (64). It was identified that different expression levels of HOXC11 represented different functions, with overexpression promoting proliferation and downregulation inhibiting proliferation (64). In addition, the study reported that high immunohistochemical expression of HOXC11 was associated with T stage, N stage and Ki67 level in RCC (65). The function of HOX genes remains unknown in oral tumorigenesis (66). However, HOXB2, HOXB7 and HOXB13 have been reported to be important in oral tissues, and dysregulation of HOXB7 results in increased tumor cell proliferation in oral squamous cell carcinoma tissues (67). Previous studies have demonstrated that HOXA1 serves an important role in tumorigenesis. It has been reported that the disorder expression of HOXA1 in several cancer types (68) is involved in regulating multiple cellular processes, including proliferation,
apoptosis, etc. (69), and is therefore closely associated with the shorter overall survival of patients after resection. In addition, HOXA1 expression was correlated with shorter overall survival in patients with small cell lung cancer (70).

HOXA1 overexpression is associated with cell growth and tumor formation in mice (60). A previous study reported that elevated HOXA1 expression levels inhibited differentiation, and resulted in transformation and the development of acute myeloblastic leukemia in lethally irradiated mice (71). Increased levels of HOXA1 in the mammary carcinoma MCF-7 cell line have also been demonstrated to dramatically increase proliferation via the signal transducers and activators of transcription pathway (72) and transcriptional upregulation of B-cell lymphoma 2 (Bcl-2) (60). HOXA1 controls the transcriptional upregulation of Bcl-2, cyclin D1 and c-Myc (73). However, various studies have reported that HOXA1 overexpression does not affect adhesion, invasion or apoptosis (60,69,71,74). The gene is dependent on the Meis, Prep or homeothorax protein cofactors to activate and/or repress transcription (75). Increased HOXA1 expression results in promoting the proliferation and decreasing the apoptosis of mammary carcinoma cells (69), thus contributing to the acceleration of cancer growth.

Previous studies have demonstrated that the upregulated expression of HOXA1 promotes proliferation, whereas downregulated HOXA1 expression decreases cell proliferation (71-73). In addition, HOXA1-positive cells correlate with N stage, T stage, proliferation and differentiation, and poor survival (76).

3. HOX genes and IncRNAs

A number of non-coding RNAs are located in and associated with HOX gene clusters, including HOXAIR1, HOX5, HOX3, HOX10, HOXB3, HOX11, HOXD3, HOX12, HOX13, HOX14, pre-B-cell leukemia HOX4, HOXA distal transcript antisense RNA (HOTTIP) and HOTAIR. The associations between downstream and upstream HOX gene clusters with IncRNAs are presented in Fig. 1.

3.1. HOTAIR. DNA-binding proteins target ncRNAs, such as HOTAIR, to specific sites in chromatin (77). HOTAIR is a
2.2-kb lncRNA implicated in the suppression of the HOX loci, which subsequently promotes breast cancer metastasis (33,78). Whether lncRNAs directly or indirectly regulate the expression of HOX gene family members has gained increasing attention in the context of development and tumorigenesis (24,78,79). Furthermore, altered HOX gene expression has been observed in certain forms of cancer (80). In a recent study, Xu et al. used a high-throughput microarray to analyze the lncRNAs and mRNA expression profiles in lung adenocarcinoma (81). A total of 564 dysregulated lncRNAs and 83 HOX subtypes (including 35 HOX genes and 48 HOX ncRNAs) were screened out, which indicated that there is a close relationship between HOX genes and lncRNAs in lung adenocarcinoma (81). Taurine upregulated 1 (TUG1) was able to modulate HOXB7 expression, which may be partially responsible for TUG1-mediated proliferation regulation, thus affecting the proliferation of non-small cell lung cancer (NSCLC) in vitro and in vivo (82). The signals of p53/TUG1/PRC2/HOXB7 may serve as targets for NSCLC diagnosis and therapy (82). A previous study observed that when HOXB7 expression was inhibited, proliferation was also subsequently inhibited (83). HOXB7 primarily regulates cell growth through the phosphoinositide 3-kinase/Akt and mitogen-activated protein kinase pathways (83). HOXB7 is able to promote cell proliferation in different types of tumors through its high levels of expression (83-85). In addition, HOXB7 has been demonstrated to be an oncogene (57,83,86,87), and its involvement is essential in various forms of cancer.

A number of studies have reported that the HOX family genes serve as classic modification targets of the polycomb complex during development, with all four gene clusters being highly enriched in histone 3 lysine 27 trimethylation (88,89). HOTAIR mediates epigenetic silencing by guiding PRC2 to the HOX loci (78). Various lncRNAs are located within the HOX gene clusters, with ~231 ncRNAs estimated to be present in humans (24). HOTAIR, a HOX lncRNA, is transcribed from the antisense of canonical HOXC genes, which interact with PRC2 and silence the subsequent transcription of the HOXD locus (78).

3.2. HOTTIP. HOTTIP is a lncRNA that targets genes through chromosomal looping (77); it is crucial for metastasis of various forms of human cancer, and its overexpression results in PRC2 silencing HOXD9, a locus involved in developmental pattern (78). A previous study observed that HOTAIR expression is increased in human breast cancer, which targets PRC2 in the genome and results in changes in gene expression (33). Overexpression of HOTAIR has also been documented in several other types of human cancer, including hepatocellular (90) and pancreatic (35) cancer, suggesting that this lncRNA has an oncogenic function. However, its potential involvement in gastric cancer remains to be elucidated. Further evidence supporting the involvement of HOTAIR in epithelial-mesenchymal transition (EMT) is the interaction between Twist and miR-10b. miR-10b is an miRNA transcribed from the 3′ HOXD locus and may be induced by Twist (36).

HOTTIP resides at the 5′ end of the HOXA locus; it has been been identified as one of the 231 ncRNAs associated with the human HOX loci (24) and has been associated with the activation of multiple 5′ HOXA genes in vivo (34). HOTTIP was previously considered as a negative prognostic factor in patients with liver cancer, and its increased expression was associated with enhanced liver cancer metastasis (38). It was observed that inhibition of HOTTIP potentiated the antitumor effects of gemcitabine in vitro and in vivo. Furthermore, HOXA13-knockdown by RNA interference (siHOXA13) demonstrated that HOTTIP promoted pancreatic ductal adenocarcinoma (PDAC) cell proliferation, invasion and chemoresistance, at least partially through regulating HOXA13 (91). In addition, immunohistochemical analysis revealed that high HOXA13 expression correlated with poor histological differentiation, lymph node metastasis and decreased overall survival in patients with PDAC (91).

A large number of HOX genes function as transcriptional regulators, and govern cell proliferation and differentiation (87). Furthermore, HOX genes have a distinctive gene network organization (four chromosomal loci: A, B, C and D) and represent the most repeat-poor regions within the human genome. The HOXA locus consists of a cluster of 11 HOX genes with a graded expression pattern along the body’s appendages from proximal (close to the main body) to distal (appendage tip) (78,88). Among the HOX genes, HOXA13, which is a marker of gut primordial posteriorization during development (78), serves an essential role in tumorigenesis of the liver and bladder, and in esophageal cancer (33,78,90). To examine the association between HOTTIP dysregulation and HOXA cluster expression, HOTTIP-knockdown studies were performed, and the expression level of the HOXA cluster (HOXA7, A9, A10, A11, and A13) was further evaluated by polymerase chain reaction (92). Although HOXA7, A9, A10, A11 and A13 gene expression was lower in the HOTTIP-knockdown group compared with the control, HOXA13 was the most significantly inhibited gene within the HOX A locus following depletion of HOTTIP in pancreatic cancer cells (92). Furthermore, it was demonstrated that the expression levels of HOTTIP and HOXA13 were significantly positively correlated with 90 PDAC tissues and paired adjacent non-neoplastic tissues (92). In addition, small interfering RNA-mediated HOXA13-knockdown inhibited the proliferation, invasion and EMT of PDAC cells, which was consistent with the functional changes that occurred after silencing the expression of HOTTIP in PDAC cells (92). Targeted inhibition of HOXA13 in two pancreatic cancer cell lines led to a decrease in the HOXA13 mRNA level and lower HOTTIP expression, which was consistent with the results of another study (93).

4. HOX gene clusters encoding miRNAs

Human HOX gene clusters exhibit a low density of interspersed repeats, and cis-regulatory elements effect HOX gene expression (94). The enhancer of the clusters is able to regulate HOXD temporal co-linearity (95). Previous studies have demonstrated that lncRNAs are involved in the regulation of HOX gene expression (96), and certain microRNAs participate in this regulation (97), including miR-196 (98) and miR-10 (99). Due to the technical limitations, earlier lncRNA studies have been performed in only two species (mouse and human), such as the previous studies of HOTAIR (33,78) and HOTAIRM1 (100). However, one study performed an integrated analyses of HOX gene clusters between the kangaroo...
family and eutherians (101). Well-studied miRNAs, including miR-196a/b, miR-10a/b and IncRNAs (such as HOTAIR, HOTAIRM1 and HOXA11-AS), were observed to dysregulate gene expression and were associated with tumor development (24,32,36). miR-464, miR-10 and miR-414 are located in the tammar HOX clusters and were identified to have an effect on these, for example inhibiting lymphangiogenesis (miR-414) and promoting active tumor cell invasion (miR-10) (102,103). In addition, certain novel miRNAs are transcribed from elsewhere on the tammar genome and regulate the expressions of HOXB and HOXD clusters by specifically interacting with the mRNAs transcribed from them (104).

5. Future prospects

A large number of studies have investigated IncRNAs in recent years, but only a few IncRNAs associated with HOX genes have been detected with sequencing and ChIP microarray. It is expected that an increasing number of HOX genes and related IncRNAs may be identified in the coming years, with further analysis required to determine their specific functions. Further studying the associations between IncRNAs and HOX will lead to a deeper understanding of the mechanism of cancer development.

Acknowledgements

This study was supported by The Key Clinical Specialty Discipline Construction Program of Fujian, The Specialized Science and Technology Key Project of Fujian Province (grant no. 2013YZZ0002-3), The Science and Technology Infrastructure Construction Program of Fujian Province (grant no. 2014Y2005), The Natural Science Foundation of Fujian Province (grant no. 2015J05174), The Scientific Research Project of Health and Family Planning Commission of Fujian Province (grant no. 2015-1-94), The Scientific Foundation of Fuzhou Health Department (grant no's. 2014-S-139-3, 2014-S-w19 and 2013-S-wp1) and the Postdoctoral Science Foundation of Fujian Infrastructure Construction Program of Fujian Province (grant no's. 2014 -S-139-3, 2015Y2005), Project of Health and Family Planning Commission of Fujian and the Postdoctoral Science Foundation of Fuzhou Health Department (grant no's. 2014 -S-139-3, 2015Y2005), The Natural Science Foundation of Fujian Province, Science and Technology Key Project of Fujian Province, The Specialized Science and Technology Project of Fujian Province (grant no. 2015J05174), The Scientific Research Project of Health and Family Planning Commission of Fujian Province (grant no. 2015-1-94), The Scientific Foundation of Fuzhou Health Department (grant no's. 2014-S-139-3, 2014-S-w19 and 2013-S-wp1) and the Postdoctoral Science Foundation of Fujian Medical University (grant no. 149318).

References


