

Hox Clusters and LncRNAs

Subjects: Biology

Contributor: Milana Kulakova

Bilaterian animals operate the clusters of Hox genes through a rich repertoire of diverse mechanisms, including, due to a large set of various non-coding RNAs. Long non-coding RNAs (lncRNAs), which are transcribed from the sense (coding) DNA strands of Hox clusters, control the work of Hox genes in the cis and trans position, are involved in the establishment and maintenance of the epigenetic code of Hox loci, and can even serve as a source of regulatory peptides, which switch cellular energy metabolism. All antisense lncRNAs in human Hox clusters are therapeutic targets for malignant tumors, and their careful study has profound practical meaning.

Keywords: long noncoding RNAs ; lncRNAs ; antisense ncRNAs ; Hox genes ; Hox clusters

1. Introduction

A little more than a century has already passed since Calvin Bridges, who worked in the laboratory of Thomas Hunt Morgan, revealed the new type of mutations in *Drosophila melanogaster*. Those mutations were localized in bithorax (bx) locus and resulted in a partial transformation of the halter to the wing.

Edward Lewis proceeded with Bridges' work and within a few decades he described, in detail, the BX-C complex, which controls the morphogenesis of thoracic and abdominal segments of flies ^[1]. This is how Hox genes were first found, the genes which are universal for all bilaterian animals, control their development and play a huge role in morphological evolution.

Hox genes certainly are the most studied developmental genes to date. Even now, there are more than 500 papers published every year investigating or discussing their functions. This undying interest in the subject can be explained by the multiple roles of Hox genes in development. These genes work through the whole embryogenesis process starting from the earliest steps of development ^[2] and until the extreme old age of multicellular animals; the vertebrates which use Hox genes to control homeostasis, being the perfect example ^{[3][4][5][6]}. The same genes specify the organization of the body plan of the animals from the largest clade—Nephrozoa. At the same time, they can be easily co-opted into the developmental programs of evolutionarily new structures, such as the photophore of fireflies ^[7] and hair follicles of mammals ^[8].

We suggest that the pervasiveness of Hox genes is caused by:

- Their fundamental role in the ground plan formation (this excludes the loss of Hox genes in most bilaterian animals);
- The simplicity of DNA-consensus for the binding of Hox homeodomain and the ability of Hox genes to form the dimers with cofactors to make this consensus more complicated;
- The complex regulation of their transcription.

Figure 1 summarizes the schematic presentations of the known mechanisms of Hox genes' transcriptional control. This complicated picture raises a number of issues. For example, it is unclear at which evolutionary step this regulatory complexity appeared. What could it be used for if it originates from the common ancestor of Nephrozoa? Which regulatory mechanisms were inherited by modern Bilateria from their common ancestor and which arose independently in different taxons? Which mechanisms are predetermined by the structure of Hox cluster? Are there any mechanisms that can underlie the clusteral structure of Hox genes? To solve these fundamental questions, the several extant monographs do not seem to be sufficient, but it is possible to come closer to an adequate understanding through analysis of separate regulatory mechanisms. In this review, we intend to discuss the accumulated data concerning long non-coding RNAs which are transcribed from the sequences of Hox clusters in the opposite direction, i.e., from the sense (coding) strand. We have focused on these regulatory transcripts because there are more of them in Hox clusters than sense lncRNAs

and because they are better studied. In addition, the regulatory potential of antisense molecules is higher due to additional options, such as the formation of duplexes with mRNAs of Hox genes.

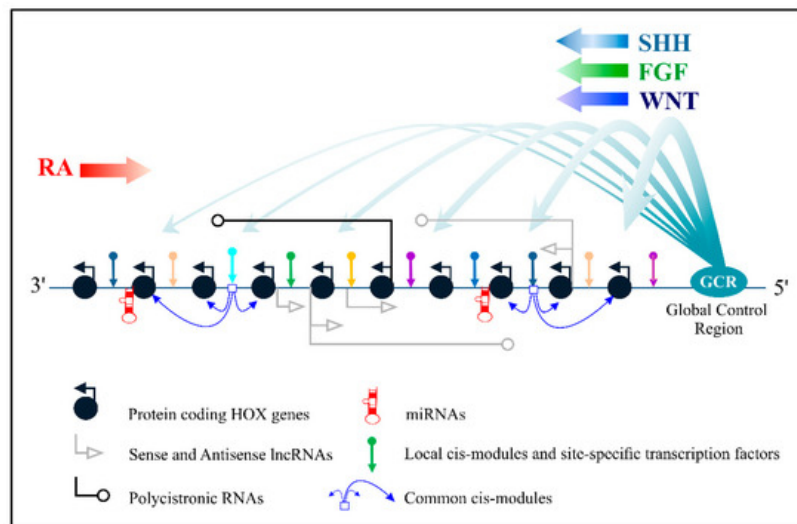


Figure 1. The main pathways for controlling the transcription of Hox genes.

Large arrows indicate Retinoic acid pathway (RA), Sonic Hedgehog pathway (SHH), WNT and FGF pathways.

The world of regulatory RNAs is huge. Since the accumulation of data on the structure and function of regulatory RNAs is actively proceeding, even their classification is not yet stable. Traditionally, non-coding RNAs are divided into short (<200 bp) and long (>200 bp) classes [8][9]. Long non-coding RNAs are the largest and the most heterogenic class of ncRNAs. These molecules were found in all the organisms studied to date including viruses. With a few exceptions, they are synthesized by RNA-polymerase II, possess the 5'-cap, polyadenylated and can exist simultaneously in both poly-A+ and poly-A- forms [10]. Their most intriguing characteristic is that in the row of multicellular animals and plants the percentage of lncRNA coding sequences in genome grows together with the increasing morphological complexity, i.e., the increasing number of cell types. For example, the sponge *Amphimedon queenslandica* (Demospongia) has no more than 15 cell types. The transcriptomic analysis of the several developmental stages of *Amphimedon* revealed 2935 lncRNAs, which is ~7.5% of protein-coding sequences (around 40,000 mRNA) of this sponge [11][12]. There slightly more than 400 cell types in the human body, around one third of which are comprised by the derivatives of the neural crest [13]. According to the various databases, the number of genes coding lncRNAs in humans varies from 140,356 (LncBook: <https://bigd.big.ac.cn/lncbook/index>) to 56,946 (LNCipedia: <https://lncipedia.org>), which, by a significant extent, exceeds the number of mRNA coding genes (20,352 according to CHES: <http://ccb.jhu.edu/chess> accessed on 15 June 2021).

It is argued that around one third of all the transcripts produced by an eukaryotic cell represent the transcriptional noise, caused by the imperfection of splicing mechanisms and mistakes in the initiation of transcription [14]. One can suggest that a huge number of lncRNA which can be found in metazoan transcriptomes is the result of these mistakes. This seems to be quite a rational view on the issue of huge lncRNA redundancy in comparison with mRNAs.

However, it is recognized that the level of tissue-specificity, i.e., the preferential transcription in the definite cell type, at least of one subclass of lncRNAs—intergenic lncRNAs (lincRNAs)—is much higher than among protein-coding genes [15]. Moreover, the large-scale analysis of transcriptional dynamics of lncRNAs in seven vertebrate species (human, rhesus macaque, mouse, rat, rabbit, opossum and chicken) revealed that during the developmental process there was a transition from the universal and conservative lncRNAs to specific and low conservative ones, which was in good accordance with the expression dynamics of protein-coding developmental genes [16].

Antisense lncRNAs are synthesized from the coding DNA strand and can overlap the sequences of protein-coding genes and genes coding other lncRNAs. In the human genome, 44,624 antisense non-coding transcripts were found (according to LNCipedia). It is generally accepted that around 70% of protein coding genes of mammals have antisense transcripts, which are synthesized from their own (independent) or divergent (bidirectional) promoters [17]. Those antisense lncRNAs that are fully or partially overlap the exons of protein-coding genes referred to as Natural Antisense Transcripts (NATs). The first evidence of the existence of the genes coding ncRNAs in Hox cluster appeared in the last century when the transcriptional activity of the bxd-region of *Drosophila* localized 5' from the *Ubx* gene was analyzed [18]. However, the complicated and variable lncRNA-dependent Hox-regulation is mostly studied in mammals and human cell cultures [19][20][21][22].

2. Antisense lncRNAs in Mammalian Hox Clusters

Four mammalian Hox clusters are composed of 39 protein-coding genes, which are transcribed in a spatially and temporally collinear manner in embryonic and definitive tissues excluding the fore- and midbrain. In their 2007 review paper, Denis Duboule classified the clusters of this type as organized [23]. These are small clusters (from 100 to 170 kb) in which all the genes are oriented in a similar fashion and as a result are transcribed in one direction. Hox clusters of mammals are free from any foreign protein-coding genes and almost free from the repeats, but can code miRNAs and lncRNAs. The fact that there are more non-protein coding transcriptional units (TUs) in Hox clusters of humans than of mice [24], and that vertebrates possess, in general, more ncRNAs in Hox clusters than protostomians (though this group is not sufficiently studied) demonstrates the functional importance and the probable contribution of these transcripts to natural selection.

The first two investigations describing the single lncRNAs of Hox genes were performed in 1995. They demonstrate the transcription of anti-*Hoxd3* [25] and anti-*Hoxa11* (now *Hoxa11os*) in mice [26]. It turned out that these molecules are polyadenylated and have several different isoforms [26]. It was shown that their expression differs from the sense transcripts of sequences they overlap [25][26]. One of the transcripts had distinct nuclear localization [25]. In both cases, antisense RNAs contained the region, which was complementary to the protein-coding part; thus, they belonged to NATs. Surprisingly, anti-*Hoxd3* transcript [25] is still absent in the database, since both possible candidate asRNAs transcribed from this locus—*Hoxd3os1* and *Gm38462*—do not possess exons overlapping homeobox and are expressed in the different parts of the embryo compared with the transcript described in the paper.

These first described cases looked like anecdotal reports. Only ten years later, the true both-way expression traffic of mammalian Hox clusters was demonstrated by using the available databases (EST and genomic databases) and new genomic technologies (Tiling array, Chip-Seq, CAGE), focused on global transcriptomic analysis [24][27][28][29]. In these large and technically complicated works, the following patterns were revealed:

First, it turned out that almost the whole of the Hox clusters of humans and mice are transcribed in both the sense and antisense directions [24][27][28]. This transcription was observed in various embryonic and definitive tissues of humans and mice, in the human cell line of teratocarcinoma [24][27] and in fibroblasts [28] from different locations of the human adult body. This last model system revealed 407 discrete transcribed regions in four Hox loci. Only 101 of them refer to the exons of Hox genes [28]. The main part of the transcribed regions (231) are the intergenic areas with three-quarters comprising antisense lncRNAs. These lncRNAs are transcribed from individual or divergent promoters. Second, there are bicystronic and polycystronic transcripts among antisense and sense (including protein-coding) RNAs [29][30]. These molecules are read from extended parts of the Hox clusters (primary transcripts up to 30 Kb), so that they overlap or include the exons of two or more genes.

It is worth noting that not only mammals but also crustaceans possess bicystronic Hox transcripts [31], as well as onychophorans and myriapods [32][33]. Their function is still unclear but it is known that RNAs of this type can be processed to form the normal mRNAs [31], and also the chimeric transcripts that contain the sequence fragments of two neighboring genes [32][33]. The presence of transcripts of this kind can consolidate the Hox cluster.

Third, antisense lncRNAs of vertebrate Hox clusters are expressed following the rule of spatial collinearity and are often coexpressed with the neighboring protein-coding genes [27][28]. Using the exogenous retinoic acid, the transcription of antisense lncRNAs was induced in the cells of embryonic teratocarcinoma (NT2D1) alongside mRNA transcription. In some cases, lncRNAs were activated even earlier than neighboring Hox genes and could serve as mediators for this induction [27]. Since the initiation of their activation in the experiment was followed by the liberation of chromatin from repression complexes, we can assume that one of the functions of persisting antisense lncRNA expression lies in the prevention of transcriptional silencing [24][27]. It is important to remember that a part of antisense lncRNAs does not co-express with the neighboring genes [28]. Moreover, RT-PCR or tiling array methods do not reveal the spatial transcription pattern in the organ or its anlage. For example, the complementary expression of *Hoxa11* and *Hoxa11os*, visualized in the limb anlage by WMISH [26], would be revealed as co-expression by the aforementioned methods.

Fourth, lncRNAs of Hox genes possess a certain evolutionary conservatism at the level of primary sequences. They are less conservative than protein-coding RNAs, but more conservative to various degrees than non-transcribed parts of the clusters [28]. The homologs of HOXB-AS3 are found in different vertebrates from the elephant shark to humans, but more importantly, there is a homology between the promoters of HOXA-AS3 and HOXB-AS3. These genes can probably be considered as ohnologs, as the protein-coding Hox genes from different clusters. This means that in the single ancestral cluster of the vertebrates, the sequence of this lncRNA already existed [34]. Moreover, a significant number of lncRNAs of

Hox genes possess the special type of evolutionary conservatism—syntenic conservatism [24][29][35][36]. In this case, the very position of lncRNAs and their functions are conserved but not their primary sequence. In terms of the evolutionary distance between *Branchiostoma lanceolatum* and *Homo sapiens*, 16 syntenic homologs were revealed [36]. The question about equivalency of these molecules' function is still open and each case needs individual analysis due to the specificity of Hox traffic regulation. We will discuss one of the examples below.

The fact that the new class of molecules contained in Hox clusters is functionally significant was indirectly confirmed by the presence of small motifs (6–8 nt), which were specific for the different molecules from spatially variable parts of the body. In other words, lncRNAs from the anterior parts of the body differ by these motifs from the transcripts of posterior parts. Motifs that were specific for proximal and distal body parts were also found [28]. Moreover, Hox clusters of vertebrates do not practically contain repeats, and the small blocks of SINE, LINE and LTR can be only found in the loci where neither sense nor antisense transcription occur [29]. However, for the full study of the functional role of lncRNAs in the processes of development and growth, experimental works on representational models are needed. It was the case that the majority of data concerning Hox lncRNA functions were obtained on a limited number of models, mainly human cell cultures. **Table 1** and **Table 2** contain the 18 most studied antisense lncRNAs from the four clusters of human Hox genes. One can see that each molecule possesses multiple functions, most of which were described for pathological processes. In most cases, these lncRNAs perform transcriptional silencing, the induction of transcription or its modulation through recruiting the proteins of chromatin remodeling and transcriptional factors to the target sites. Here, lncRNAs work as scaffolds for the assembling of the protein complexes, both repressing and activating. Moreover, most of them work as competing endogenous RNAs (ceRNAs), which scavenge many miRNAs of different types as a sponge. Almost all of these molecules influence Hox gene expression in cis or trans. Hox genes themselves also regulate the transcription lncRNAs. For example, it was shown that HOXB13 directly interact with the promotor and induce the transcription of HOXC-AS3, which further provoke proliferation, migration and invasion of glioblastoma (GBM) cells [37].

Table 1. Antisense lncRNAs from four human Hox clusters. The zones of the overlap of lncRNA exons and RNA exons from the opposite (template) strand are shown in red.

LncRNA	Length (nt)	Type of LncRNA	Position in the Hox Cluster
HOXA Cluster			
<i>HOTAIRM1</i>	4000; 1052; 783	Linc	<i>HOXA1 – HOXA2</i>
<i>HOXA-AS2</i>	1048	Linc or NAT	<i>HOXA3 – HOXA4</i>
<i>HOXA-AS3</i>	3918; 3992	Linc or NAT	<i>HOXA4 – A5, – A6 – HOXA7</i>
<i>HOXA10-AS</i>	1161	Linc or NAT	<i>HOXA9 – HOXA10</i>
<i>HOXA11-AS</i>	1628	Linc or NAT	<i>HOXA11 – HOXA13</i>
<i>HOTTIP</i>	4665	Linc or NAT	<i>HOXA13 – EVX1</i>
HOXB Cluster			
<i>HOXB-AS1</i>	797	Linc or NAT	<i>HOXB2 – HOXB3</i>
<i>HOXB-AS2</i>	3594	NAT RNA	<i>HOXB3</i>
<i>HOXB-AS3</i>	785; 611; 549; 545; 514; 452; 446; 336	Linc or NAT *	<i>HOXB4 – B5 – HOXB6</i>
<i>HOXB-AS4</i>	543; 513	Linc	<i>HOXB9 – HOXA13</i>
<i>PRAC2</i>	1193; 560; 518; 503; 448	Linc *	<i>HOXB9 – HOXA13</i>
HOXC Cluster			
<i>HOXC-AS1</i>	548	Intronic	<i>HOXC9</i>
<i>HOXC-AS2</i>	504	Linc or NAT	<i>HOXC9 – HOXC10</i>
<i>HOXC-AS3</i>	368	Linc or NAT	<i>HOXC10 – HOXC11</i>
<i>HOTAIR</i>	2370; 2364; 2337	Linc	<i>HOXC11 – HOXC12</i>

LncRNA	Length (nt)	Type of LncRNA	Position in the Hox Cluster
<i>HOXC13-AS</i>	1408	Linc or NAT	<i>HOXC9</i> →
HOXD Cluster			
<i>HAGLR</i>	4086; 4037; 4007; 3942; 3923; 3905; 3893; 3891; 3821; 3812; 3794; 3782	Linc or NAT	<i>HOXD1</i> – <i>HOXD3</i>
<i>HOXD-AS2</i>	692	Linc or NAT	<i>HOXD4</i> → <i>D8</i> → <i>D9</i> → <i>HOXD10</i>

LncRNAs of Hox clusters are transcribed directly (sense) and in an opposite way (antisense). Though the number of antisense transcripts is considerably larger, there is no impression that they function according to different rules. In fact, the transcripts of both types may overlap the large parts of Hox clusters, may be intronic and intergenic, may work in cis- and trans-mode, possess conservatism to a greater or lesser degree and effectively function as scaffolds. However, the important regulatory mechanisms exist, which can be only performed through the antisense transcripts. In **Table 1**, the majority of lncRNAs are marked as “linc or NAT” because, among the registered isoforms of one RNA, both types were found. The large part of these molecules indeed have exons that partially overlap with the exons of Hox genes. Theoretically, this means that NATs can interact with mRNAs of Hox genes and either stabilize or cleave them to siRNA, thus participating in post-transcriptional silencing. Moreover, the antisense lncRNAs can repress the transcription of sense transcripts through the mechanism of interference at the level of transcriptional complexes [38]. This mechanism is realized in the cell nucleus.

Table 2. Antisense lncRNAs from four human Hox clusters and their main functions and targets.

As ncRNA	Functions	Mechanism of Work	Localization	Targets	Orthologs	Discovered in	Refs
HOTAIRM1	Control of the cell cycle in the myeloid cell lineage; control of the differentiation of granulocytes; control of neuronal differentiation timing; control of osteogenesis in dental follicle stem cells	Serve as protein scaffolds; Enhancer; Sponges big set miRNAs	Nucleus Cytoplasm	HOXA cluster; NEUROGENIN 2; miR-196b; miR-125b	Chordata	2009	[36] [39] [40] [41] [42] [43] [44]
HOXA-AS2	Promotion of proliferation, migration and invasion in many types of tumors; regulation EMT; negative regulates endothelium inflammation	Serve as protein scaffolds; Sponges big set miRNAs	Nucleus Cytoplasm	c-MYC; EGFR; Bax/TRAIL; EZH2/LSD1; PBX3; NF-kB; miR-373	Primates	2013	[45] [46] [47] [48] [49]
HOXA-AS3	Control of cell cycle, proliferation, migration and apoptosis in many types of cancer cells; positive regulation of endothelium inflammation; activation the MEK/ERK Signaling Pathway	Stabilization of HOXA6 mRNA, sponges miR-29c and mir-455-5p	Nucleus Cytoplasm	HOXA6; NF-kB; miR-29c	Homo sapiens	2017	[50] [51] [52] [53] [54]

As ncRNA	Functions	Mechanism of Work	Localization	Targets	Orthologs	Discovered in	Refs
HOXA10-AS	Cell cycle and apoptosis control in glioma, lung adenocarcinoma (LAD), oral cancer and acute myeloid leukemia (AML) cells	?	Cytoplasm	HOXA10; Wnt pathway; NF-kB	Birds Mammals	2018	[55] [56] [57] [58]
HOXA11-AS	Control of the menstrual cycle; cell cycle, proliferation, migration and apoptosis control in many types of cancer cells	Serve as protein scaffolds; Sponges big set miRNAs	Nucleus Cytoplasm	HOXA11; TGF- β pathway, LATS1; CyclinD1; CyclinE; CDK4; CDK2	Rodents Primates Bamboo shark	2002	[59]
HOTTIP	Control of 5'HOXA genes' transcription during development. Participation in pathogenesis of almost all types of cancer.	Activates HOXA genes through recruiting of WDR5 and MLL	Nucleus	HOXA7-HOXA13; LSD1; EZH2; IL-6; miR-30b	Rodents Primates Bamboo shark	2011	[60] [61]
HOXB-AS1	Glioblastoma and endometrial carcinoma and multiple myeloma (MM) promotion	ILF3-mediated activation of HOXB3 and HOXB3 transcription; stabilization of their mRNAs; stabilization of FUT4 mRNA	Nucleus Cytoplasm	HOXB2; HOXB3; Wnt pathway; FUT4; miR-186-5p; miR-149-3p	Homo sapiens	2019	[62] [63] [64]
HOXB-AS2	Potentially participate in the development of atrial fibrillation	?	?	?	Homo sapiens	2020	[65]
HOXB-AS3	Control of energetic metabolism in the cell through alternating the isoforms of pyruvate kinase M (PKM); promoting of the cancer processes through repression of p53 transcription; activation of PI3/AKT pathway	Codes the conservative peptide of 53 amino acids long, which is important for PKM splicing; Sponges miRNAs	Nucleus Cytoplasm	PKM; DNMT1; p53; I3K-AKT-mTOR pathway; miR-378a-3p	Homo sapiens Rodents	2017	[66] [67] [68] [69] [70]
HOXB-AS4	The sequence is differentially methylated in normal and pancreatic cancer cells	?	?	?	Homo sapiens	2018	[71]
HOXB-AS5 or PRAC2	Associated with breast cancer (lncRNA) and prostate cancer (protein)	Encodes 140 aa nuclear protein	Nucleus (protein)	I3K-AKT-mTOR pathway (lncRNA)	Artiodactyla Bats Colugo Primates	2003 (protein) 2017 (lncRNA)	[72] [73]

As ncRNA	Functions	Mechanism of Work	Localization	Targets	Orthologs	Discovered in	Refs
HOXC-AS1	Cholesterol homeostasis participation, inhibition of atherosclerosis; promotion of growth and metastatic formation in several types of malignant tumors	Promotes the transcription and translation of HOXC6; boosts c-MYC mRNA	Nucleus Cytoplasm	HOXC6; miR-590-3p; c-MYC; Wnt pathway; miR-590-3p	Homo sapiens	2016	[74] [75] [76] [77] [78]
HOXC-AS2	Promotion of growth and metastatic formation in several types of malignant tumors	Promote HOXC13 transcription; can sponge miR-876-5p to affect ZEB1 expression	Nucleus Cytoplasm	HOXC13; ZEB1; miR-876-5p	Homo sapiens	2019	[79] [80] [81] [82]
HOXC-AS3	Functions under the direct control of HOXB1. Promotion of growth and metastatic formation in several types of malignant tumors	Promote the transcription of 5'HOXC genes; stabilizes HOXC10 mRNA; can sponge miR-3922-5p, impairs the maturation of miR-96	Nucleus	HOXC8; HOXC9; HOXC10; HOXC11; HOXC12; HOXC13; YBX1; thymidine kinase 1 (TK1); FOXM1; miR-96; miR-3922-5p	Homo sapiens	2018	[37] [81] [82] [83] [84] [85] [86] [87] [88] [89]
HOTAIR	Reprogramming of chromatin state to promote cancer metastasis; PRC2 and PRC2-independent induction of transcriptional repression. Promotion of growth and metastatic formation in several types of malignant tumors	Scaffold: A 5' domain of HOTAIR binds PRC2, whereas a 3' domain of HOTAIR binds the LSD1/CoREST/REST complex; Sponging big set microRNA	Nucleus Cytoplasm	HOXD cluster (40 Kb in 5'area) HOXA1; HOXA5; HOXC11; p53; p27; E-cadherin; NOTCH1/JAGGED1; SNAIL; GLI2; Protocadherin 10; Wnt pathway; Dozens of miRNAs, critical for proliferation and differentiation control	Rodents Carnivores Primates Marsupials	2007	[28] [35] [90] [91] [92] [93] [94] [95] [96] [97] [98] [99] [100] [101] [102] [103] [104] [105] [106] [107] [108] [109] [110] [111] [112] [113] [114] [115] [116] [117] [118] [119]
HOXC13-AS	Promotion of proliferation, migration and invasion of cells of several cancer types	Sponging big set microRNA	Cytoplasm	HOXC13; c-MYC; miR-383-3p; miR-497-5p	Homo sapiens	2019	[120] [121] [122] [123] [124]

As ncRNA	Functions	Mechanism of Work	Localization	Targets	Orthologs	Discovered in	Refs
HAGLR or HOXD-AS1	Impairment of HAGLR regulation (up or down) to promote growth and metastatic formation of several types of malignant tumors	Binds to WDR5 and EZH2 for activation and repression of target genes	Nucleus Cytoplasm	HOXD3; JAK2/STAT3 pathway; Wnt pathway; Ras/ERK pathway; TGF- β pathway; p57; miR-133a-3p; miR-133b; miR-130a; miR-185-5p	Rodents Primates	2014	[125]
							[126]
							[127]
							[128]
							[129]
							[130]
							[131]
							[132]
							[133]
							[134]
							[135]
							[136]
							[137]
							[138]
HOXD-AS2	Downregulation of HOXD-AS2 significantly promotes the progression of gastric cancer	?	Cytoplasm	HOXD8; PI3K/Akt pathway	Homo sapiens	2018	[139]
							[140]
							[141]
							[142]
							[143]
							[144]
							[145]
							[146]
							[147]
							[148]
							[149]

Surprisingly, not all lncRNAs can be positively considered as non-protein coding. The small (less than 100 codons in length) open reading frames (small open reading frames; smORFs) are found in many lncRNAs. Their functional coding potential was under debate for a long time, but the peptides were later found in mammals; these were synthesized from “non-coding” templates. Around a dozen and a half short molecules, from 9 to 250 amino acids long, are synthesized from the templates of precursors of miRNAs, lncRNAs and even from one transcribed repeat [150][151]. The functional meaning of some of the peptides from lncRNAs of Hox clusters was shown in experiments. HOXB-AS3 lncRNA codes a peptide that is 53 amino acids long. It turned out that this peptide functions as a switch in the energetic metabolism of the cell, since it controls the splicing of one of the most important catabolic enzyme—pyruvate kinase M (PKM). In the presence of this peptide, the isoform M1 of kinase is synthesized, which shifts the energetic metabolism towards the oxidative phosphorylation. In the absence of the peptide, another isoform of the pyruvate kinase M—M2—prevails in the cell, which promotes the glycolysis. The first isoform dominates in the most differentiating somatic cells, while the second one is needed in the actively proliferating stem cells that use anaerobic glycolysis as a source of “fast” ATP synthesis. The majority of human tumors demonstrate the reduced transcription of HOXB-AS3 and start to synthesize PKM2. The metabolism of the tumor cell is forced to undergo glycolysis since the mitochondria, the membranes of which are used for oxidative phosphorylation, are not functional [66]. Importantly, Hox genes do not function in the actively proliferating ESM, and the synthesis of HOXB-AS3 alongside the regulatory peptide occurs later, when the intensity of embryonic cell proliferation goes down.

We would like to discuss the high evolutionary and regulatory plasticity of Hox lncRNAs using the examples of the two most studied molecules. In 2007, the first human antisense lincRNA—HOTAIR—was discovered and functionally characterized [28]. It is transcribed from the region between HOXC11 and HOXC12 and functions as a scaffold for assembling of the protein complexes PRC2 (Polycomb Repressive Complex 2) and LSD1 (Lysine-Specific Demethylase 1). Due to these two complexes, HOTAIR realizes two functions at the same time—the methylation of H3K27 and demethylation of H3K4me2 [90]. HOTAIR starts to work early in human embryogenesis and is needed for transcriptional silencing of the 5'-area of HOXD cluster [28]. Moreover, it controls the work of a large group of miRNAs which are critical for development and differentiation. Dozens of transcriptional factors and components of signaling pathways function under the direct or indirect control of HOTAIR. The defects in regulation of HOTAIR transcription were found in many types of malignant tumors. Surprisingly, this functional load does not bring about evolutionary conservatism. The similarity between human and mouse HOTAIR comprises only 55% [91]. Besides, TUs, which are positionally equivalent to HOTAIR, are transcribed from the homological cluster regions of many mammals including kangaroos [92]. Contrary to the human homolog, murine Hotair does not contain the motifs for interaction with the proteins of chromatin remodeling and cannot participate in cis- or trans-silencing, which was demonstrated by the deletion of a large fragment of the HoxC locus [93]. The absence of Hotair does not critically interfere with the development of mice. The mutational effects did not exceed the natural variability. Therefore, Denis Duboule et al. have termed the effects caused by Hotair mutations “homeopathic rather than homeotic” [94]. However, if the deletion of the certain nucleotides of Hotair is performed, the impairment of the

transcription of neighboring protein-coding Hox genes occurs also due to the formation of new lncRNAs—“ghostair” and “antiHotair”—which form as a result of this deletion. Consequently, the punctual deletion of Hotair in mice can provoke the formation of a homeotic phenotype that is close to the expected one, taking into account the conservative function of human and murine Hox orthologues [20][94]. Thus, the functional conservatism of lncRNAs cannot be estimated through the mutant phenotype without detailed analysis of mechanisms underlying this phenotype.

Another fascinating example of the functional plasticity of lncRNA in one organism is an antisense lincRNA HOTAIRM1 (HOXA Transcript Antisense RNA, Myeloid-Specific 1). The HOTAIRM1 gene is localized between HOXA1 and HOXA2. This lncRNA is conservative at the level of its position in the cluster. Its syntenic homologs are found in birds, amphibians, bony fishes and even in lancelet [36]. The global studies on the role of Hotairm1 in the mammalian development were not yet performed but it was shown that it controls the differentiation of myelopoietic cells and osteogenesis in the descendants of dental follicle stem cells [19][39]. In the cell line NT2-D1 (embryonal pluripotent carcinoma), Hox gene transcription can be induced by retinoic acid (RA). In this cell line, HOXA1 and HOXA2 are located topologically close to HOXA4, HOXA5 and HOXA6 due to chromatin looping before the RA induction. After the induction with RA, the transcription of all Hox genes is initiated, but HOTAIRM1 modulates their work in such a manner that they all become collinear at the transcriptional level [152].

The synthesis of two different isoforms of HOTAIRM1 containing three exons—the long non-spliced and the short spliced ones—is also initiated by RA. The first isoform recruits the protein complex UTX/MLL (H3K27-demethylase/H3K4 methyltransferase) that initiates euchromatinization of HOXA1 and HOXA2 loci and, thus, their transcription. The short HOTAIRM1 isoform interacts with PRC2 and lowers the transcription level of HOXA4, HOXA5 and HOXA6. This process is followed by reassociation of the proximal and distal parts of the cluster, but the initial spatial vicinity between HOTAIRM1 gene and other proximal genes allows the short repressing isoforms to reach the target sites by diffusion. This truly concise and elegant mechanism is not realized in another cell line NB4 (acute promyelocytic leukemia) with a different chromatin landscape. HOTAIRM1 isoform from NB4 cells contains two but not three exons and does not control the proximal genes [152]. It was recently discovered that HOTAIRM1 represses the transcription of NEUROGENIN 2 in Hox-independent manner and, thus, is localized at the top of regulatory cascade, which regulates the timing of neuronal differentiation [40].

Thus, we can safely assume that lncRNAs of mammalian Hox clusters are numerous, multifunctional and changeable (**Table 2**). At the same time, they can be observed in the same regulatory continuum with conservative Hox genes and both participate in the reciprocal regulation [24][28][37][152]. In this connection, taking into account the large set of functions that are critical for tissue homeostasis (and probably for development), it is still an enigma as to how lncRNAs can demonstrate such a high evolutionary plasticity. The most provocative aspect here is the structural and functional divergence of the human and murine HOTAIR.

We suggest that this evolutionary paradox can be solved if observed from the ontogenetic point of view. The multiple and variable functions of lncRNAs may be easily delegated. They can be duplicated in the organism by other components of GRNs and, in case of their dysregulation, this impairment can be compensated during the developmental process. Local GRNs that control the differentiation of multi-, bi and unipotential stem cells of the adult organism do not possess the compensatory potential of the complicated multilevel embryonic GRNs. This can explain the participation of Hox lncRNAs and Hox genes themselves in the tumor formation.

Thus, in the hyperconservative compact mammalian Hox clusters, the molecular machinery of antisense ncRNAs is internalized, constantly generating changes in their regulation. The range of this variation is defined by the stabilizing selection that happens also at the level of the cell populations in the developing embryo. This is a breeding ground for ontogenetic and evolutionary variability.

3. Antisense lncRNAs in Hox Clusters of Protostomian Animals

The increasing interest in lncRNAs raised the number of studies in which different insects, including those that did not belong to the classical models of molecular and genetic research, were used [50][83][153][154][155][156][157]. From these studies, it has emerged that insects possess lesser amounts than vertebrates but still significant repertoires of these molecules (from 2949 for *Anopheles gambiae* to 11,810 for *Bombyx mori*).

There is no surprise that *Drosophila* is the best studied protostomian animal in the sense of genome structure and molecular architecture of development. It is worth noting that the first genes that were mapped in the BX-C complex in the

classical paper of E. Lewis in 1978 ^[1] belonged mostly not to the protein-coding sequences but to the regulatory elements that, nevertheless, were transcribed ^{[18][139]}.

In a research study performed in 2002, a series of 1–2 Kb Dig-probes was synthesized. These probes overlapped the intergenic region between *abd-A* and *abd-B* genes of *Drosophila*. This intergenic region of 100 kb length was called *iab* (infraabdominal) and contained cis-regulatory elements that control the nearby Hox genes ^[139]. Almost all the probes used revealed the transcription of unknown lncRNAs in fly embryos. This transcription preceded Hox mRNA expression and demonstrated spatial collinearity that was similar to mRNA patterns.

Only three probes revealed antisense transcription and, at present, we can identify them with two ncRNAs presented in FlyBase (<http://flybase.org>): Dmel\lncRNA:iab4 (CR31271) and Dmel\lncRNA:CR43617. The first transcript belongs to the specific type of lncRNAs that are processed to become miRNA (mir-iab4). Mature mir-iab4 takes the *Ubx* gene under negative control ^{[158][159]}. The biological activity of lncRNA iab-4 seems to be limited by this function. The mir-iab4 sequence is evolutionarily conserved and can be traced in insects, chelicerates and crustaceans ^[160]. The function of the second transcript remains unclear.

In the modern database FlyBase (<http://flybase.org>) dozens of lncRNAs are presented, including those which are produced from ANTP-C and BX-C regions, but their functions are still unknown. It is worth mentioning that there are no NATs among these non-coding transcripts, but intronic and intergenic antisense lncRNAs are highly presented. Since the Hox cluster of *D. melanogaster* as well as of other drosophilids is disrupted and contains microinversions, it cannot be regarded as an ancestral cluster of insects ^[161]. One can assume that the repertoire of Hox-associated lncRNAs is depleted in the fly but is preserved in other species, for example, in the intact Hox cluster of the beetle *Tribolium castaneum*.

The work of Shippy et al. (2008) describes the transcriptional activity of the beetle Hox cluster studied by tiling arrays ^[162]. It turned out that the Hox cluster of *T. Castaneum* produced multiple non-coding transcripts between 0 and 72 developmental hours. However, the authors described only two antisense lncRNAs, one of which is transcribed from the intergenic region between *Antp* and *Ubx* genes, and the second—from the intron of *Ubx* gene. Thus, despite the huge amount of data, the NATs of Hox genes of insects either do not exist or they are not yet found.

Two research studies were focused on lncRNAs in crustaceans; notably, in both cases, single-molecule real-time (SMRT) sequencing by PacBio was used. A total of 3958 lncRNAs were found in pacific white shrimp (*Litopenaeus vannamei*) ^[163]. This fact demonstrates the essential similarity of hexapods and crustaceans in the number of lncRNAs. Wan et al. (2019) managed to sequence 23,644 long non-coding RNAs in crab *Scylla paramamosain* (Decapoda). However, at the time of the paper's publication, the genome assembling was not yet finished. The authors admitted that they were not able to assess the false rate of identified lncRNAs without the fully assembled genomic data ^[164].

One gets the impression that despite a significant number of the assembled genomes in the Pancrustacea clade being of high quality, there is still the lack of data about the functions and even the number of lncRNAs. It is worth noting that it is the representatives of crustaceans who possess the most compact Hox clusters among all arthropods—192.8 Kb in *Paracyclopina nana* (Copepod) and 324.6 Kb in *Daphnia magna* (Branchiopoda) ^{[165][166]}. Here, as well as in the case of vertebrates, the mechanisms consolidating the cluster exist but are not yet studied.

Myriapoda is a sister group to Pancrustacea. The genomes of the following three myriapods are assembled to date: *Strigamia maritima*, *Helicorhombus holstii* and *Trigoniulus corallines*. These belong to the two large sister branches—Centipede and Millipede ^{[167][168]}. The authors did not focus on the analysis of lncRNAs but clarified the genomic organization of myriapod Hox clusters and revealed their important ancestral feature. It turned out that Hox clusters of *Strigamia* (Centipede) and *Trigoniulus* (Millipede) are flanked with the orthologs of *Eve-skipped* (*eve*) gene on their 5'-ends. The traces of this ancient synteny were earlier found in chordates and some cnidarian species. This feature, together with the compactness of Hox clusters (457 Kb from *labial* to *eve* in *S. maritima*), make the myriapods a prospective model for studying the mechanisms of Hox-regulation, including those that involve NATs. Myriapods were the first Protostomia for which Hox-associated NATs were found and then studied ^{[32][33]}. NAT of *Ubx* (*aUbx*) of *Strigamia* is cloned (GenBank: DQ368689.1), occupies the region of 1051 nt and is complementary to 3'UTR of *Ubx* mRNA (GenBank: DQ368688.1) within sequences that are at least 226 bases long. *aUbx* promotor is localized 3' from the sequence of *Ubx* gene and closer to 5'-end of *Antp* gene. *aUbx* starts to express earlier and more anterior than its sense transcript, simultaneously with *Antp* mRNA. In addition, sense and antisense *Ubx* patterns look mutually exclusive ^[32].

In two different myriapod species—the millipede *Glomeris marginata* and the centipede *Lithobius forficatus*—*aUbx* RNAs were revealed with sense probes. The expression patterns of antisense RNAs were complementary to *Ubx* mRNAs,

which indicates the high conservation of this regulatory relationship in myriapod clade [33]. Outside myriapods (in onychophora *Euperipatoides* and spider *Cupiennius*), no *aUbx* or any other NATs transcribed from Hox loci were found, at least by sense probes [33].

Among spiralian animals—the second protostomian clade—antisense lncRNAs of Hox genes were found in nereidid annelids [169][170]. Using sense probes, our working group showed the presence of NATs of *Hox5* and *Hox7* in the postlarval development and regeneration of *Alitta virens*. Later, we revealed the antisense transcription for almost all genes of *A. virens* and *Platynereis dumerilii* Hox clusters [171]. We managed to clone some of these NATs. The structure of the *P. dumerilii* Hox cluster is not yet published. The *A. virens* Hox cluster is studied by pulsed-field gel electrophoresis [172]. According to these data, the cluster is not atomized and does not exceed 2.4 Mb in length. In general, we observe a rich repertoire of antisense patterns for nereidid Hox genes, some of which are complementary to mRNA patterns and some demonstrate the large overlapping zones between sense and antisense areas of transcription. NATs of Hox genes are revealed in the developing segmented larva, in growing and regenerating juvenile worms and, presumably, are present at all stages of the worm's life cycle. The functional role of these transcripts is unclear and needs to be studied in further experiments. From the analysis of the expression of *Avi-antiHox5*, it is obvious that the antisense transcript is quickly up-regulated at the amputation site while *Hox5* mRNA gradually vanishes from this area [170].

In general, despite the huge amount of genomic data and detailed analysis of whole genome transcriptional dynamics performed for model and non-model protostomian animals, little attention is focused on regulatory RNAs, particularly NATs. The only hope is that “big data” for the most transcriptomes studied by the modern methods already contain the needed information, which should be included in the analysis.

4. The Implications for the Uprise of Antisense LncRNAs in Hox Clusters and the Reasons for Their Evolutionary Maintenance

Thus, Hox-associated antisense lncRNAs were found in mammals, insects, myriapods and nereidid polychaetes (Figure 2). This regulatory principle may have already existed in the common ancestor of Nephrozoa. Nevertheless, it does not appear that those molecules were found in all model systems where they were looked for. There are probably some specific “demands” of natural selection on the functioning of the Hox cluster which support—or, on the contrary, alleviate—the priority of this regulative mode.

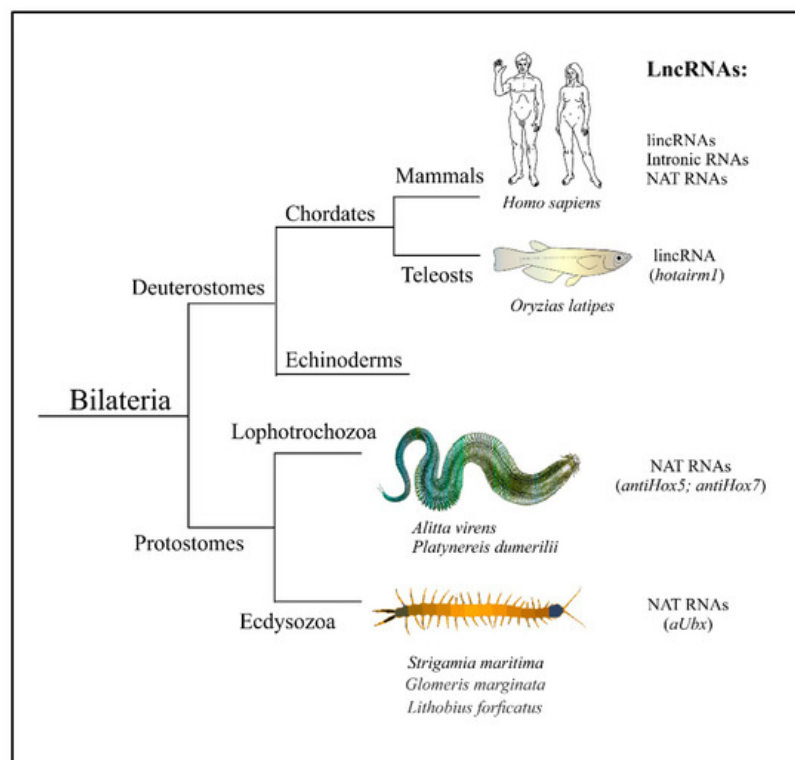


Figure 2. Hox-associated lncRNAs are found in bilateral animals belonging to three main evolutionary lineages—Deuterostomia, Lophotrochozoa, and Ecdysozoa. Individual human lncRNAs are listed in Table 1.

We can assume that one of the reasons for the large number of Hox-associated antisense lncRNAs (as well as sense ones) in mammals and other vertebrates is the necessity for dose compensation of Hox-ohnologs. This compensation can be realized due to the delicate epigenetic tuning, where the mediators between Hox sequences and remodeling proteins

or transcriptional factors are lncRNAs working in both cis- and trans-positions. However, this function cannot be primary and the only one since, among the animals possessing one Hox cluster, the profound antisense transcription was also unraveled (myriapods, annelids).

The complicated regulation of gene functioning demands the extensional regulatory sequences, but it is not true for mammals. There are probably restrictions of the Hox cluster size caused by the specificity of their early activation in vertebrate development. During gastrulation, the consecutive liberation of Hox loci from heterochromatized territories occur from 3' to 5' ends of the clusters and, thus, the sequential temporal activation of Hox genes is manifested [173]. This process is directly connected to the speed of primary mesenchymal cells' (PMCs) ingression through the primary stick [174]. The more 5'-Hox genes are activated in the ingressing cells, the slower they are internalized and the more posterior is the location they obtain in the end. By these means, the spatial Hox code of vertebrates is realized through the temporal dynamics of euchromatization, which, in turn, depends on the physical size of Hox locus. An abundance of intergenic and polycistronic lncRNAs (sense and antisense) in Hox clusters may function as mediators in the process of early euchromatization and stabilize the local epigenetic states. They can also keep the cluster from relaxation and disruptions.

It is unknown whether the compact cluster of vertebrate ancestors allowed them to use the system of lncRNAs, or whether these molecules themselves possessed sufficient functional importance to resist the natural selection pressure and keep the Hox cluster from relaxation and reorganization. To establish the cause and effect in this scenario, the lncRNAs' loss-of-function experiments in vertebrates and comparative data concerning Hox traffic in the species with variably organized clusters are needed. At this point, it is worth mentioning that animals with disorganized clusters that include a gap (for example, *Drosophila*) possess very poor repertoires of lncRNAs in comparison with vertebrates. In the lancelet, which preserved the integral cluster, the syntenic homolog of *Hotairm1* was found even though it is absent in tunicates and appendicularians with disorganized and atomized clusters [36]. Interestingly, *Hotairm1* was not found even in cyclostomes with an integral but relaxed Hox cluster [36][175].

Recent studies have revealed a fascinating detail. It turned out that antisense transcription is not an exotic mode of regulation but a natural means of maintaining the chromatin of functional genes in a dynamic state. By the dynamic state, the authors mean the high level of histone turnover, i.e., the circulation of histones in the promoter and gene body [176]. It was shown in yeast studies (*S. cerevisiae*) that antisense transcription produces dynamic chromatin [176]. If we imagine that the gene has a final number of configurations at the chromatin level and each of these configurations corresponds to a certain level of expression, then the dynamic chromatin allows the gene to switch between these configurations [176]. This can be the reason why the genes that change their expression level in response to stress or environmental changes most commonly possess the antisense transcripts. Moreover, antisense transcription initiated by bidirectional promoters can spread the regulatory signals from one locus to the neighboring genes [177]. There is one more remarkable detail. The antisense transcription represses poorly activated sense transcription but if the sense transcription is strong enough, the antisense transcription vanishes itself [177]. This means that the sense transcript will only be transcribed in the presence of the threshold amount of the activator. This, in turn, predetermines the differential gene activation in the morphogen gradient. Genes comprised into well-ordered clusters can easily utilize these basic characters for co-regulation.

From this perspective, we can view the antisense transcription in Hox clusters from a different angle. Mammals and nereidid annelids use Hox genes for their whole life for development, growth, and reparative and physiological regeneration. This means that they have a constant demand for dynamic chromatin in Hox loci. Centipede *Strigamia* forms a finite number of segments in embryogenesis (epimorphic type of development). However, the basal myriapods form the terminal number of segments only after a few moults have passed (hemianamorphic type of development). Thus, they also need the postembryonic work of Hox genes [178]. One can probably expect that antisense Hox lncRNAs will be found in those animals that retain the integral clusters and keep using Hox genes in postembryonic life.

References

1. Lewis, E.B. A gene complex controlling segmentation in *Drosophila*. *Nature* 1978, 7, 565–570.
2. Maslakov, G.P.; Kulishkin, N.S.; Surkova, A.A.; Kulakova, M.A. Maternal Transcripts of Hox Genes are Found in Oocytes of *Platynereis dumerilii* (Annelida, Nereididae). *J. Dev. Biol.* 2021. Unpublished work.
3. Awgulewitsch, A. Hox in hair growth and development. *Naturwissenschaften* 2003, 90, 193–211.
4. Brun, A.C.; Björnsson, J.M.; Magnusson, M.; Larsson, N.; Leveén, P.; Ehinger, M.; Nilsson, E.; Karlsson, S. *Hoxb4*-deficient mice undergo normal hematopoietic development but exhibit a mild proliferation defect in hematopoietic stem cells. *Blood* 2004, 103, 4126–4133.

5. Argiropoulos, B.; Humphries, R.K. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* 2007, 26, 6766–6776.
6. Morgan, R.; Whiting, K. Differential expression of HOX genes upon activation of leukocyte sub-populations. *Int. J. Hematol.* 2008, 87, 246–249.
7. Stansbury, M.S.; Moczek, A.P. The function of Hox and appendage-patterning genes in the development of an evolutionary novelty, the *Photuris* firefly lantern. *Proc. Biol. Sci.* 2014, 281, 20133333.
8. Horabin, J.I. Long noncoding RNAs as metazoan developmental regulators. *Chromosome Res.* 2013, 21, 673–684.
9. Brosnan, C.A.; Voinnet, O. The long and the short of noncoding RNAs. *Curr. Opin. Cell Biol.* 2009, 21, 416–425.
10. Ponting, C.P.; Belgard, T.G. Transcribed dark matter: Meaning or myth? *Hum. Mol. Genet.* 2010, 19, R162–R168.
11. Fernandez-Valverde, S.L.; Calcino, A.D.; Degnan, B.M. Deep developmental transcriptome sequencing uncovers numerous new genes and enhances gene annotation in the sponge *Amphimedon queenslandica*. *BMC Genom.* 2015, 16, 387.
12. Gaiti, F.; Fernandez-Valverde, S.L.; Nakanishi, N.; Calcino, A.D.; Yanai, I.; Tanurdzic, M.; Degnan, B.M. Dynamic and Widespread lncRNA Expression in a Sponge and the Origin of Animal Complexity. *Mol. Biol. Evol.* 2015, 32, 2367–2382.
13. Vickaryous, M.K.; Hall, B.K. Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biol. Rev. Camb. Philos. Soc.* 2006, 81, 425–455.
14. Pertea, M.; Shumate, A.; Pertea, G.; Varabyou, A.; Breitwieser, F.P.; Chang, Y.-C.; Madugundu, A.K.; Pandey, A.; Salzberg, S.L. CHES: A new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. *Genome Biol.* 2018, 19, 208.
15. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 2011, 25, 1915–1927.
16. Sarropoulos, I.; Marin, R.; Cardoso-Moreira, M.; Kaessmann, H. Developmental dynamics of lncRNAs across mammalian organs and species. *Nature* 2019, 571, 510–514.
17. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium; et al. Antisense transcription in the mammalian transcriptome. *Science* 2005, 309, 1564–1566.
18. Lipshitz, H.D.; Peattie, D.A.; Hogness, D.S. Novel transcripts from the Ultrabithorax domain of the bithorax complex. *Genes Dev.* 1987, 1, 307–322.
19. Zhang, X.; Lian, Z.; Padden, C.; Gerstein, M.B.; Rozowsky, J.; Snyder, M.; Gingeras, T.R.; Kapranov, P.; Weissman, S.M.; Newburger, P.E. A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. *Blood* 2009, 113, 2526–2534.
20. Li, L.; Liu, B.; Wapinski, O.L.; Tsai, M.C.; Qu, K.; Zhang, J.; Carlson, J.C.; Lin, M.; Fang, F.; Gupta, R.A.; et al. Targeted disruption of Hota1r leads to homeotic transformation and gene derepression. *Cell Rep.* 2013, 5, 3–12.
21. Zhao, H.; Zhang, X.; Frazão, J.B.; Condino-Neto, A.; Newburger, P.E. HOX antisense lincRNA HOXA-AS2 is an apoptosis repressor in all trans retinoic acid treated NB4 promyelocytic leukemia cells. *J. Cell Biochem.* 2013, 114, 2375–2383.
22. Pradeepa, M.M.; McKenna, F.; Taylor, G.C.A.; Bengani, H.; Grimes, G.R.; Wood, A.J.; Bhatia, B.; Bickmore, W.A. Psp1/p52 regulates posterior Hoxa genes through activation of lncRNA Hottip. *PLoS Genet.* 2017, 13, e1006677.
23. Duboule, D. The rise and fall of Hox gene clusters. *Development* 2007, 134, 2549–2560.
24. Engström, P.G.; Suzuki, H.; Ninomiya, N.; Akalin, A.; Sessa, L.; Lavorgna, G.; Brozzi, A.; Luzi, L.; Tan, S.L.; Yang, L.; et al. Complex Loci in human and mouse genomes. *PLoS Genet.* 2006, 2, e47.
25. Bedford, M.; Arman, E.; Orr-Urtreger, A.; Lonai, P. Analysis of the Hoxd-3 gene: Structure and localization of its sense and natural antisense transcripts. *DNA Cell Biol.* 1995, 14, 295–304.
26. Hsieh-Li, H.M.; Witte, D.P.; Weinstein, M.; Branford, W.; Li, H.; Small, K.; Potter, S.S. Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility. *Development* 1995, 121, 1373–1385.
27. Sessa, L.; Breiling, A.; Lavorgna, G.; Silvestri, L.; Casari, G.; Orlando, V. Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. *RNA* 2007, 13, 223–239.
28. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Bruggmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 2007, 129, 1311–1323.

29. Mainguy, G.; Koster, J.; Woltering, J.; Jansen, H.; Durston, A. Extensive polycistronism and antisense transcription in the mammalian Hox clusters. *PLoS ONE* 2007, 2, e356.
30. Simeone, A.; Pannese, M.; Acampora, D.; D'Esposito, M.; Boncinelli, E. At least three human homeoboxes on chromosome 12 belong to the same transcription unit. *Nucleic Acids Res.* 1988, 16, 5379–5390.
31. Shiga, Y.; Sagawa, K.; Takai, R.; Sakaguchi, H.; Yamagata, H.; Hayashi, S. Transcriptional readthrough of Hox genes *Ubx* and *Antp* and their divergent post-transcriptional control during crustacean evolution. *Evol. Dev.* 2006, 8, 407–414.
32. Brena, C.; Chipman, A.D.; Minelli, A.; Akam, M. Expression of trunk Hox genes in the centipede *Strigamia maritima*: Sense and anti-sense transcripts. *Evol. Dev.* 2006, 8, 252–265.
33. Janssen, R.; Budd, G.E. Gene expression suggests conserved aspects of Hox gene regulation in arthropods and provides additional support for monophyletic Myriapoda. *Evodevo* 2010, 1, 4.
34. Degani, N.; Ainbinder, E.; Ulitsky, I. Highly conserved and cis-acting lncRNAs produced from paralogous regions in the center of HOXA and HOXB clusters in the endoderm lineage. *bioRxiv* 2020.
35. Diederichs, S. The four dimensions of noncoding RNA conservation. *Trends Genet.* 2014, 30, 121–123.
36. Herrera-Úbeda, C.; Marín-Barba, M.; Navas-Pérez, E.; Gravemeyer, J.; Albuixech-Crespo, B.; Wheeler, G.N.; Garcia-Fernández, J. Microsyntenic Clusters Reveal Conservation of lncRNAs in Chordates Despite Absence of Sequence Conservation. *Biology* 2019, 8, 61.
37. Wang, X.; Sun, Y.; Xu, T.; Qian, K.; Huang, B.; Zhang, K.; Song, Z.; Qian, T.; Shi, J.; Li, L. HOXB13 promotes proliferation, migration, and invasion of glioblastoma through transcriptional upregulation of lncRNA HOXC-AS3. *J. Cell Biochem.* 2019, 120, 15527–15537.
38. Latgé, G.; Poulet, C.; Bours, V.; Josse, C.; Jerusalem, G. Natural Antisense Transcripts: Molecular Mechanisms and Implications in Breast Cancers. *Int. J. Mol. Sci.* 2018, 19, 123.
39. Chen, Z.; Zheng, J.; Hong, H.; Chen, D.; Deng, L.; Zhang, X.; Ling, J.; Wu, L. lncRNA HOTAIRM1 promotes osteogenesis of hDFSCs by epigenetically regulating HOXA2 via DNMT1 in vitro. *J. Cell Physiol.* 2020, 235, 8507–8519.
40. Rea, J.; Menci, V.; Tollis, P.; Santini, T.; Armaos, A.; Garone, M.G.; Iberite, F.; Cipriano, A.; Tartaglia, G.G.; Rosa, A.; et al. HOTAIRM1 regulates neuronal differentiation by modulating NEUROGENIN 2 and the downstream neurogenic cascade. *Cell Death Dis.* 2020, 11, 527.
41. Zhang, X.; Weissman, S.M.; Newburger, P.E. Long intergenic non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid maturation in NB4 human promyelocytic leukemia cells. *RNA Biol.* 2014, 11, 777–787.
42. Díaz-Beyá, M.; Brunet, S.; Nomdedéu, J.; Pratcorona, M.; Cordeiro, A.; Gallardo, D.; Escoda, L.; Tormo, M.; Heras, I.; Ribera, J.M.; et al. The lincRNA HOTAIRM1, located in the HOXA genomic region, is expressed in acute myeloid leukemia, impacts prognosis in patients in the intermediate-risk cytogenetic category, and is associated with a distinctive microRNA signature. *Oncotarget* 2015, 6, 31613–31627.
43. Shi, T.; Guo, D.; Xu, H.; Su, G.; Chen, J.; Zhao, Z.; Shi, J.; Wedemeyer, M.; Attenello, F.; Zhang, L.; et al. HOTAIRM1, an enhancer lncRNA, promotes glioma proliferation by regulating long-range chromatin interactions within HOXA cluster genes. *Mol. Biol. Rep.* 2020, 47, 2723–2733.
44. Li, Y.Q.; Sun, N.; Zhang, C.S.; Li, N.; Wu, B.; Zhang, J.L. Inactivation of lncRNA HOTAIRM1 caused by histone methyltransferase RIZ1 accelerated the proliferation and invasion of liver cancer. *Eur. Rev. Med. Pharmacol. Sci.* 2020, 24, 8767–8777.
45. Fang, Y.; Wang, J.; Wu, F.; Song, Y.; Zhao, S.; Zhang, Q. Long non-coding RNA HOXA-AS2 promotes proliferation and invasion of breast cancer by acting as a miR-520c-3p sponge. *Oncotarget* 2017, 8, 46090–46103.
46. Zhang, Y.; Xu, J.; Zhang, S.; An, J.; Zhang, J.; Huang, J.; Jin, Y. HOXA-AS2 Promotes Proliferation and Induces Epithelial-Mesenchymal Transition via the miR-520c-3p/GPC3 Axis in Hepatocellular Carcinoma. *Cell Physiol. Biochem.* 2018, 50, 2124–2138.
47. Zhu, X.; Liu, Y.; Yu, J.; Du, J.; Guo, R.; Feng, Y.; Zhong, G.; Jiang, Y.; Lin, J. LncRNA HOXA-AS2 represses endothelium inflammation by regulating the activity of NF-κB signaling. *Atherosclerosis* 2019, 281, 38–46.
48. Xiao, S.; Song, B. LncRNA HOXA-AS2 promotes the progression of prostate cancer via targeting miR-509-3p/PBX3 axis. *Biosci. Rep.* 2020, 40, BSR20193287.
49. Wang, J.; Su, Z.; Lu, S.; Fu, W.; Liu, Z.; Jiang, X.; Tai, S. LncRNA HOXA-AS2 and its molecular mechanisms in human cancer. *Clin. Chim. Acta* 2018, 485, 229–233.
50. Wu, F.; Zhang, C.; Cai, J.; Yang, F.; Liang, T.; Yan, X.; Wang, H.; Wang, W.; Chen, J.; Jiang, T. Upregulation of long noncoding RNA HOXA-AS3 promotes tumor progression and predicts poor prognosis in glioma. *Oncotarget* 2017, 8,

51. Zhang, H.; Liu, Y.; Yan, L.; Zhang, M.; Yu, X.; Du, W.; Wang, S.; Li, Q.; Chen, H.; Zhang, Y.; et al. Increased levels of the long noncoding RNA, HOXA-AS3, promote proliferation of A549 cells. *Cell Death Dis.* 2018, 9, 707.
52. Tong, Y.; Wang, M.; Dai, Y.; Bao, D.; Zhang, J.; Pan, H. LncRNA HOXA-AS3 Sponges miR-29c to Facilitate Cell Proliferation, Metastasis, and EMT Process and Activate the MEK/ERK Signaling Pathway in Hepatocellular Carcinoma. *Hum. Gene Ther. Clin. Dev.* 2019, 30, 129–141.
53. Zhu, X.; Chen, D.; Liu, Y.; Yu, J.; Qiao, L.; Lin, S.; Chen, D.; Zhong, G.; Lu, X.; Wang, Y.; et al. Long Noncoding RNA HOXA-AS3 Integrates NF- κ B Signaling to Regulate Endothelium Inflammation. *Mol. Cell Biol.* 2019, 39, e00139-19.
54. Chen, W.; Li, Q.; Zhang, G.; Wang, H.; Zhu, Z.; Chen, L. LncRNA HOXA-AS3 promotes the malignancy of glioblastoma through regulating miR-455-5p/USP3 axis. *J. Cell Mol. Med.* 2020, 24, 11755–11767.
55. Dong, C.Y.; Cui, J.; Li, D.H.; Li, Q.; Hong, X.Y. HOXA10-AS: A novel oncogenic long non-coding RNA in glioma. *Oncol. Rep.* 2018, 40, 2573–2583.
56. Sheng, K.; Lu, J.; Zhao, H. ELK1-induced upregulation of lncRNA HOXA10-AS promotes lung adenocarcinoma progression by increasing Wnt/ β -catenin signaling. *Biochem. Biophys. Res. Commun.* 2018, 501, 612–618.
57. Al-Kersh, S.; Bhayadia, R.; Ng, M.; Verboon, L.; Emmrich, S.; Gack, L.; Schwarzer, A.; Strowig, T.; Heckl, D.; Klusmann, J.H. The stem cell-specific long noncoding RNA HOXA10-AS in the pathogenesis of KMT2A-rearranged leukemia. *Blood Adv.* 2019, 3, 4252–4263.
58. Yan, X.; Cong, B.; Chen, Q.; Liu, L.; Luan, X.; Du, J.; Cao, M. Silencing lncRNA HOXA10-AS decreases cell proliferation of oral cancer and HOXA10-antisense RNA can serve as a novel prognostic predictor. *J. Int. Med. Res.* 2020, 48.
59. Wei, C.; Zhao, L.; Liang, H.; Zhen, Y.; Han, L. Recent advances in unraveling the molecular mechanisms and functions of HOXA11-AS in human cancers and other diseases (Review). *Oncol. Rep.* 2020, 43, 1737–1754.
60. Wang, K.C.; Yang, Y.W.; Liu, B.; Sanyal, A.; Corces-Zimmerman, R.; Chen, Y.; Lajoie, B.R.; Protacio, A.; Flynn, R.A.; Gupta, R.A.; et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 2011, 472, 120–124.
61. Ghafouri-Fard, S.; Dashti, S.; Taheri, M. The HOTTIP (HOXA transcript at the distal tip) lncRNA: Review of oncogenic roles in human. *Biomed. Pharmacother.* 2020, 127, 110158.
62. Chen, R.; Zhang, X.; Wang, C. LncRNA HOXB-AS1 promotes cell growth in multiple myeloma via FUT4 mRNA stability by ELAVL1. *J. Cell Biochem.* 2020, 121, 4043–4051.
63. Bi, Y.; Mao, Y.; Su, Z.; Du, J.; Ye, L.; Xu, F. HOXB-AS1 accelerates the tumorigenesis of glioblastoma via modulation of HOBX2 and HOBX3 at transcriptional and posttranscriptional levels. *J. Cell Physiol.* 2021, 236, 93–106.
64. Liu, D.; Qiu, M.; Jiang, L.; Liu, K. Long Noncoding RNA HOXB-AS1 Is Upregulated in Endometrial Carcinoma and Sponged miR-149-3p to Upregulate Wnt10b. *Technol. Cancer Res. Treat.* 2020, 19.
65. Shi, X.; Shao, X.; Liu, B.; Lv, M.; Pandey, P.; Guo, C.; Zhang, R.; Zhang, Y. Genome-wide screening of functional long noncoding RNAs in the epicardial adipose tissues of atrial fibrillation. *Biochim. Biophys. Acta Mol. Basis Dis.* 2020, 1866, 165757.
66. Huang, J.Z.; Chen, M.; Chen, D.; Gao, X.C.; Zhu, S.; Huang, H.; Hu, M.; Zhu, H.; Yan, G.R. A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth. *Mol. Cell.* 2017, 68, 171–184.
67. Zhang, X.M.; Chen, H.; Zhou, B.; Zhang, Q.Y.; Liao, Y.; Wang, J.S.; Wang, Z.H. lncRNA HOXB-AS3 promotes hepatoma by inhibiting p53 expression. *Eur. Rev. Med. Pharmacol. Sci.* 2018, 22, 6784–6792.
68. Jiang, W.; Kai, J.; Li, D.; Wei, Z.; Wang, Y.; Wang, W. lncRNA HOXB-AS3 exacerbates proliferation, migration, and invasion of lung cancer via activating the PI3K-AKT pathway. *J. Cell Physiol.* 2020, 235, 7194–7203.
69. Huang, H.H.; Chen, F.Y.; Chou, W.C.; Hou, H.A.; Ko, B.S.; Lin, C.T.; Tang, J.L.; Li, C.C.; Yao, M.; Tsay, W.; et al. Long non-coding RNA HOXB-AS3 promotes myeloid cell proliferation and its higher expression is an adverse prognostic marker in patients with acute myeloid leukemia and myelodysplastic syndrome. *BMC Cancer* 2019, 19, 617.
70. Xu, S.; Jia, G.; Zhang, H.; Wang, L.; Cong, Y.; Lv, M.; Xu, J.; Ruan, H.; Jia, X.; Xu, P.; et al. lncRNA HOXB-AS3 promotes growth, invasion and migration of epithelial ovarian cancer by altering glycolysis. *Life Sci.* 2021, 264, 118636.
71. Ishihara, H.; Yamashita, S.; Amano, R.; Kimura, K.; Hirakawa, K.; Ueda, T.; Murakami, Y.; Tamori, A.; Tanabe, K.; Kawada, N.; et al. Pancreatic Cancer Cell Fraction Estimation in a DNA Sample. *Oncology* 2018, 95, 370–379.
72. Olsson, P.; Motege, A.; Bera, T.K.; Lee, B.; Pastan, I. PRAC2: A new gene expressed in human prostate and prostate cancer. *Prostate* 2003, 56, 123–130.

73. Rui, J.; Chunming, Z.; Binbin, G.; Na, S.; Shengxi, W.; Wei, S. IL-22 promotes the progression of breast cancer through regulating HOXB-AS5. *Oncotarget* 2017, 8, 103601–103612.
74. Huang, C.; Hu, Y.W.; Zhao, J.J.; Ma, X.; Zhang, Y.; Guo, F.X.; Kang, C.M.; Lu, J.B.; Xiu, J.C.; Sha, Y.H.; et al. Long Noncoding RNA HOXC-AS1 Suppresses Ox-LDL-Induced Cholesterol Accumulation Through Promoting HOXC6 Expression in THP-1 Macrophages. *DNA Cell Biol.* 2016, 35, 722–729.
75. Dong, Y.; Li, X.; Lin, Z.; Zou, W.; Liu, Y.; Qian, H.; Jia, J. HOXC-AS1-MYC regulatory loop contributes to the growth and metastasis in gastric cancer. *J. Exp. Clin. Cancer Res.* 2019, 38, 502.
76. Takayama, K.I.; Fujimura, T.; Suzuki, Y.; Inoue, S. Identification of long non-coding RNAs in advanced prostate cancer associated with androgen receptor splicing factors. *Commun. Biol.* 2020, 3, 393.
77. Zhou, C.; An, N.; Cao, C.; Wang, G. lncRNA HOXC-AS1 promotes gastric cancer via binding eIF4AIII by activating Wnt/ β -catenin signaling. *J. Gene Med.* 2020, 22, e3202.
78. Zhang, S.; Wang, L.; Gao, Y.; Fan, Y.; Zhang, G.; Zhang, Y. Molecular Mechanism of 73HOXC-AS1-Activated Wnt β -Catenin Signaling and eIF4AIII in Promoting Progression of Gastric Cancer. *Biomed Res. Int.* 2021, 2021, 8814843.
79. Dong, N.; Guo, J.; Han, S.; Bao, L.; Diao, Y.; Lin, Z. Positive feedback loop of lncRNA HOXC-AS2/miR-876-5p/ZEB1 to regulate EMT in glioma. *Onco Targets Ther.* 2019, 12, 7601–7609.
80. Turner, D.C.; Gorski, P.P.; Maasar, M.F.; Seaborne, R.A.; Baumert, P.; Brown, A.D.; Kitchen, M.O.; Erskine, R.M.; Dos-Remedios, I.; Voisin, S.; et al. DNA methylation across the genome in aged human skeletal muscle tissue and muscle-derived cells: The role of HOX genes and physical activity. *Sci. Rep.* 2020, 10, 15360.
81. Fu, T.; Ji, X.; Bu, Z.; Zhang, J.; Wu, X.; Zong, X.; Fan, B.; Jia, Z.; Ji, J. Identification of key long non-coding RNAs in gastric adenocarcinoma. *Cancer Biomark.* 2020, 27, 541–553.
82. Liu, B.; Li, J.; Li, J.M.; Liu, G.Y.; Wang, Y.S. HOXC-AS2 mediates the proliferation, apoptosis, and migration of non-small cell lung cancer by combining with HOXC13 gene. *Cell Cycle* 2021, 20, 236–246.
83. Yang, Z.; Hu, T. Long noncoding RNA HOXC-AS3 facilitates the progression of invasive mucinous adenocarcinomas of the lung via modulating FUS/FOX1. *Vitr. Cell Dev. Biol. Anim.* 2020, 56, 15–23.
84. Zhang, E.; He, X.; Zhang, C.; Su, J.; Lu, X.; Si, X.; Chen, J.; Yin, D.; Han, L.; De, W. A novel long noncoding RNA HOXC-AS3 mediates tumorigenesis of gastric cancer by binding to YBX1. *Genome Biol.* 2018, 19, 154.
85. Li, B.; Han, H.; Song, S.; Fan, G.; Xu, H.; Zhou, W.; Qiu, Y.; Qian, C.; Wang, Y.; Yuan, Z.; et al. HOXC10 Regulates Osteogenesis of Mesenchymal Stromal Cells Through Interaction with Its Natural Antisense Transcript lncHOXC-AS3. *Stem Cells* 2019, 37, 247–256.
86. Shi, S.H.; Jiang, J.; Zhang, W.; Sun, L.; Li, X.J.; Li, C.; Ge, Q.D.; Zhuang, Z.G. A Novel lncRNA HOXC-AS3 Acts as a miR-3922-5p Sponge to Promote Breast Cancer Metastasis. *Cancer Investig.* 2020, 38, 1–12.
87. Enteghami, M.; Ghorbani, M.; Zamani, M.; Galehdari, H. HOXC10 is significantly overexpressed in colorectal cancer. *Biomed. Rep.* 2020, 13, 18.
88. Su, J.; Yu, B.; Zhang, C.; Yi, P.; Li, H.; Xu, C.; Cao, L.; Chen, P.; Li, M.; Shen, K.; et al. Long noncoding RNA HOXC-AS3 indicates a poor prognosis and regulates tumorigenesis by binding to YBX1 in breast cancer. *Am. J. Transl. Res.* 2020, 12, 6335–6350.
89. Yang, B.; Sun, L.; Liang, L. lncRNA HOXC-AS3 Suppresses the Formation of Mature miR-96 in Ovarian Cancer Cells to Promote Cell Proliferation. *Reprod. Sci.* 2021.
90. Spitale, R.C.; Tsai, M.C.; Chang, H.Y. RNA templating the epigenome: Long noncoding RNAs as molecular scaffolds. *Epigenetics* 2011, 6, 539–543.
91. Battistelli, C.; Cicchini, C.; Santangelo, L.; Tramontano, A.; Grassi, L.; Gonzalez, F.J.; de Nonno, V.; Grassi, G.; Amicone, L.; Tripodi, M. The Snail repressor recruits EZH2 to specific genomic sites through the enrollment of the lncRNA HOTAIR in epithelial-to-mesenchymal transition. *Oncogene* 2017, 36, 942–955.
92. He, S.; Liu, S.; Zhu, H. The sequence, structure and evolutionary features of HOTAIR in mammals. *BMC Evol. Biol.* 2011, 11, 102.
93. Schorderet, P.; Duboule, D. Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genet.* 2011, 7, e1002071.
94. Amândio, A.R.; Necsulea, A.; Joye, E.; Mascrez, B.; Duboule, D. Hotair Is Dispensable for Mouse Development. *PLoS Genet.* 2016, 12, e1006232.
95. Khalil, A.M.; Guttman, M.; Huarte, M.; Garber, M.; Raj, A.; Morales, D.R.; Thomas, K.; Presser, A.; Bernstein, B.E.; van Oudenaarden, A.; et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* 2009, 106, 11667–11672.

96. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010, 329, 689–693.
97. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 2010, 464, 1071–1076.
98. Hung, T.; Chang, H.Y. Long noncoding RNA in genome regulation: Prospects and mechanisms. *RNA Biol.* 2010, 7, 582–585.
99. Schiavo, G.; D'Antò, V.; Cantile, M.; Procino, A.; Di Giovanni, S.; Valletta, R.; Terracciano, L.; Baumhoer, D.; Jundt, G.; Cillo, C. Deregulated HOX genes in ameloblastomas are located in physical contiguity to keratin genes. *J. Cell Biochem.* 2011, 112, 3206–3215.
100. Wutz, A. RNA-mediated silencing mechanisms in mammalian cells. *Prog. Mol. Biol. Transl. Sci.* 2011, 101, 351–376.
101. Guil, S.; Soler, M.; Portela, A.; Carrère, J.; Fonalleras, E.; Gómez, A.; Villanueva, A.; Esteller, M. Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat. Struct. Mol. Biol.* 2012, 19, 664–670.
102. Wu, Y.; Zhang, L.; Wang, Y.; Li, H.; Ren, X.; Wei, F.; Yu, W.; Wang, X.; Zhang, L.; Yu, J.; et al. Long noncoding RNA HOTAIR involvement in cancer. *Tumour Biol.* 2014, 35, 9531–9538.
103. Deng, Q.; Sun, H.; He, B.; Pan, Y.; Gao, T.; Chen, J.; Ying, H.; Liu, X.; Wang, F.; Xu, Y.; et al. Prognostic value of long non-coding RNA HOTAIR in various cancers. *PLoS ONE* 2014, 9, e110059.
104. Zhou, X.; Chen, J.; Tang, W. The molecular mechanism of HOTAIR in tumorigenesis, metastasis, and drug resistance. *Acta Biochim. Biophys. Sin.* 2014, 46, 1011–1015.
105. Dong, C.; Liu, S.; Lv, Y.; Zhang, C.; Gao, H.; Tan, L.; Wang, H. Long non-coding RNA HOTAIR regulates proliferation and invasion via activating Notch signalling pathway in retinoblastoma. *J. Biosci.* 2016, 41, 677–687.
106. Kalwa, M.; Hänzelmann, S.; Otto, S.; Kuo, C.C.; Franzen, J.; Joussen, S.; Fernandez-Rebollo, E.; Rath, B.; Koch, C.; Hofmann, A.; et al. The lncRNA HOTAIR impacts on mesenchymal stem cells via triple helix formation. *Nucleic Acids Res.* 2016, 44, 10631–10643.
107. Pawłowska, E.; Szczepanska, J.; Blasiak, J. The Long Noncoding RNA HOTAIR in Breast Cancer: Does Autophagy Play a Role? *Int. J. Mol. Sci.* 2017, 18, 2317.
108. Zhang, Z.; Zhu, H.; Liu, Y.; Quan, F.; Zhang, X.; Yu, L. LncRNA HOTAIR mediates TGF- β 2-induced cell growth and epithelial-mesenchymal transition in human lens epithelial cells. *Acta Biochim. Biophys. Sin.* 2018, 50, 1028–1037.
109. Botti, G.; Scognamiglio, G.; Aquino, G.; Liguori, G.; Cantile, M. LncRNA HOTAIR in Tumor Microenvironment: What Role? *Int. J. Mol. Sci.* 2019, 20, 2279.
110. Shen, J.J.; Zhang, C.H.; Chen, Z.W.; Wang, Z.X.; Yang, D.C.; Zhang, F.L.; Feng, K.H. LncRNA HOTAIR inhibited osteogenic differentiation of BMSCs by regulating Wnt/ β -catenin pathway. *Eur. Rev. Med. Pharmacol. Sci.* 2019, 23, 7232–7246.
111. Bai, J.Y.; Jin, B.; Ma, J.B.; Liu, T.J.; Yang, C.; Chong, Y.; Wang, X.; He, D.; Guo, P. HOTAIR and androgen receptor synergistically increase GLI2 transcription to promote tumor angiogenesis and cancer stemness in renal cell carcinoma. *Cancer Lett.* 2021, 498, 70–79.
112. Zhang, J.; Qiu, W.Q.; Zhu, H.; Liu, H.; Sun, J.H.; Chen, Y.; Shen, H.; Qian, C.L.; Shen, Z.Y. HOTAIR contributes to the carcinogenesis of gastric cancer via modulating cellular and exosomal miRNAs level. *Cell Death Dis.* 2020, 11, 780.
113. Ghafouri-Fard, S.; Dashti, S.; Farsi, M.; Taheri, M. HOX transcript antisense RNA: An oncogenic lncRNA in diverse malignancies. *Exp. Mol. Pathol.* 2021, 118, 104578.
114. Zhou, Y.H.; Cui, Y.H.; Wang, T.; Luo, Y. Long non-coding RNA HOTAIR in cervical cancer: Molecular marker, mechanistic insight, and therapeutic target. *Adv. Clin. Chem.* 2020, 97, 117–140.
115. Ren, M.M.; Xu, S.; Wei, Y.B.; Yang, J.J.; Yang, Y.N.; Sun, S.S.; Li, Y.J.; Wang, P.Y.; Xie, S.Y. Roles of HOTAIR in lung cancer susceptibility and prognosis. *Mol. Genet. Genomic Med.* 2020, 8, e1299.
116. Mozdarani, H.; Ezzatizadeh, V.; Rahbar Parvaneh, R. The emerging role of the long non-coding RNA HOTAIR in breast cancer development and treatment. *J. Transl. Med.* 2020, 18, 152.
117. Battistelli, C.; Garbo, S.; Riccioni, V.; Montaldo, C.; Santangelo, L.; Vandelli, A.; Strippoli, R.; Tartaglia, G.G.; Tripodi, M.; Cicchini, C. Design and Functional Validation of a Mutant Variant of the LncRNA HOTAIR to Counteract Snail Function in Epithelial-to-Mesenchymal Transition. *Cancer Res.* 2021, 81, 103–113.
118. Cantile, M.; Di Bonito, M.; De Bellis, M.T.; Botti, G. Functional Interaction among lncRNA HOTAIR and MicroRNAs in Cancer and Other Human Diseases. *Cancers* 2021, 13, 570.

119. Liang, Q.; Asila, A.; Deng, Y.; Liao, J.; Liu, Z.; Fang, R. Osteopontin-induced lncRNA HOTAIR expression is involved in osteoarthritis by regulating cell proliferation. *BMC Geriatr.* 2021, 21, 57.
120. Gao, C.; Lu, W.; Lou, W.; Wang, L.; Xu, Q. Long noncoding RNA HOXC13-AS positively affects cell proliferation and invasion in nasopharyngeal carcinoma via modulating miR-383-3p/HMGA2 axis. *J. Cell Physiol.* 2019, 234, 12809–12820.
121. Li, X.; Wang, Q.; Rui, Y.; Zhang, C.; Wang, W.; Gu, J.; Tang, J.; Ding, Y. HOXC13-AS promotes breast cancer cell growth through regulating miR-497-5p/PTEN axis. *J. Cell Physiol.* 2019, 234, 22343–22351.
122. Liu, N.; Wang, Z.; Liu, D.; Xie, P. HOXC13-AS-miR-122-5p-SATB1-C-Myc feedback loop promotes migration, invasion and EMT process in glioma. *Onco Targets Ther.* 2019, 12, 7165–7173.
123. Zhou, J.F.; Shi, Y.T.; Wang, H.G.; Yang, X.Z.; Wu, S.N. Overexpression of long noncoding RNA HOXC13-AS and its prognostic significance in hepatocellular carcinoma. *Eur. Rev. Med. Pharmacol. Sci.* 2019, 23, 7369–7374.
124. Li, W.; Zhu, Q.; Zhang, S.; Liu, L.; Zhang, H.; Zhu, D. HOXC13-AS accelerates cell proliferation and migration in oral squamous cell carcinoma via miR-378g/HOXC13 axis. *Oral Oncol.* 2020, 111, 104946.
125. Yarmishyn, A.A.; Batagov, A.O.; Tan, J.Z.; Sundaram, G.M.; Sampath, P.; Kuznetsov, V.A.; Kurochkin, I.V. HOXD-AS1 is a novel lncRNA encoded in HOXD cluster and a marker of neuroblastoma progression revealed via integrative analysis of noncoding transcriptome. *BMC Genom.* 2014, 15, S7.
126. Li, J.; Zhuang, C.; Liu, Y.; Chen, M.; Chen, Y.; Chen, Z.; He, A.; Lin, J.; Zhan, Y.; Liu, L.; et al. Synthetic tetracycline-controllable shRNA targeting long non-coding RNA HOXD-AS1 inhibits the progression of bladder cancer. *J. Exp. Clin. Cancer Res.* 2016, 35, 99.
127. Lu, C.; Ma, J.; Cai, D. Increased HAGLR expression promotes non-small cell lung cancer proliferation and invasion via enhanced de novo lipogenesis. *Tumour Biol.* 2017, 39.
128. Zheng, L.; Chen, J.; Zhou, Z.; He, Z. Knockdown of long non-coding RNA HOXD-AS1 inhibits gastric cancer cell growth via inactivating the JAK2/STAT3 pathway. *Tumour Biol.* 2017, 39.
129. Zhang, Y.; Dun, Y.; Zhou, S.; Huang, X.H. lncRNA HOXD-AS1 promotes epithelial ovarian cancer cells proliferation and invasion by targeting miR-133a-3p and activating Wnt/ β -catenin signaling pathway. *Biomed. Pharmacother.* 2017, 96, 1216–1221.
130. Gu, P.; Chen, X.; Xie, R.; Han, J.; Xie, W.; Wang, B.; Dong, W.; Chen, C.; Yang, M.; Jiang, J.; et al. lncRNA HOXD-AS1 Regulates Proliferation and Chemo-Resistance of Castration-Resistant Prostate Cancer via Recruiting WDR5. *Mol. Ther.* 2017, 25, 1959–1973.
131. Hu, Y.C.; Wang, A.M.; Lu, J.K.; Cen, R.; Liu, L.L. Long noncoding RNA HOXD-AS1 regulates proliferation of cervical cancer cells by activating Ras/ERK signaling pathway. *Eur. Rev. Med. Pharmacol. Sci.* 2017, 21, 5049–5055.
132. Chen, Y.; Zhao, F.; Cui, D.; Jiang, R.; Chen, J.; Huang, Q.; Shi, J. HOXD-AS1/miR-130a sponge regulates glioma development by targeting E2F8. *Int. J. Cancer* 2018, 142, 2313–2322.
133. Gu, W.; Zhang, E.; Song, L.; Tu, L.; Wang, Z.; Tian, F.; Aikenmu, K.; Chu, G.; Zhao, J. Long noncoding RNA HOXD-AS1 aggravates osteosarcoma carcinogenesis through epigenetically inhibiting p53 via EZH2. *Biomed. Pharmacother.* 2018, 106, 890–895.
134. Li, L.; Wang, Y.; Zhang, X.; Huang, Q.; Diao, Y.; Yin, H.; Liu, H. Long non-coding RNA HOXD-AS1 in cancer. *Clin. Chim. Acta* 2018, 487, 197–201.
135. Xia, H.; Jing, H.; Li, Y.; Lv, X. Long noncoding RNA HOXD-AS1 promotes non-small cell lung cancer migration and invasion through regulating miR-133b/MMP9 axis. *Biomed. Pharmacother.* 2018, 106, 156–162.
136. Yang, M.H.; Zhao, L.; Wang, L.; Ou-Yang, W.; Hu, S.S.; Li, W.L.; Ai, M.L.; Wang, Y.Q.; Han, Y.; Li, T.T.; et al. Nuclear lncRNA HOXD-AS1 suppresses colorectal carcinoma growth and metastasis via inhibiting HOXD3-induced integrin β 3 transcriptional activating and MAPK/AKT signalling. *Mol. Cancer* 2019, 18, 31.
137. Sun, W.; Nie, W.; Wang, Z.; Zhang, H.; Li, Y.; Fang, X. lnc HAGLR Promotes Colon Cancer Progression Through Sponging miR-185-5p and Activating CDK4 and CDK6 in vitro and in vivo. *Onco Targets Ther.* 2020, 13, 5913–5925.
138. Pan, L.X.; Ding, W. lncRNA HAGLR accelerates femoral neck fracture healing through negatively regulating miRNA-19a-3p. *Eur. Rev. Med. Pharmacol. Sci.* 2020, 24, 4080–4087.
139. Bae, E.; Calhoun, V.C.; Levine, M.; Lewis, E.B.; Drewell, R.A. Characterization of the intergenic RNA profile at abdominal-A and Abdominal-B in the *Drosophila* bithorax complex. *Proc. Natl. Acad. Sci. USA* 2002, 99, 16847–16852.
140. Qi, Y.; Wang, Z.; Wu, F.; Yin, B.; Jiang, T.; Qiang, B.; Yuan, J.; Han, W.; Peng, X. Long noncoding RNA HOXD-AS2 regulates cell cycle to promote glioma progression. *J. Cell Biochem.* 2018.

141. Yao, L.; Ye, P.C.; Tan, W.; Luo, Y.J.; Xiang, W.P.; Liu, Z.L.; Fu, Z.M.; Lu, F.; Tang, L.H.; Xiao, J.W. Decreased expression of the long non-coding RNA HOXD-AS2 promotes gastric cancer progression by targeting HOXD8 and activating PI3K/Akt signaling pathway. *World J. Gastrointest. Oncol.* 2020, 12, 1237–1254.
142. Galis, F.; Van Dooren, T.J.; Feuth, J.D.; Metz, J.A.; Witkam, A.; Ruinard, S.; Steigenga, M.J.; Wijnaendts, L.C. Extreme selection in humans against homeotic transformations of cervical vertebrae. *Evolution* 2006, 60, 2643–2654.
143. Yang, L.; Wang, Y.; Lu, Y.; Li, B.; Chen, K.; Li, C. Genome-wide Identification and Characterization of Long Non-coding RNAs in *Tribolium castaneum*. *Insect Sci.* 2020.
144. Wu, Y.; Cheng, T.; Liu, C.; Liu, D.; Zhang, Q.; Long, R.; Zhao, P.; Xia, Q. Systematic Identification and Characterization of Long Non-Coding RNAs in the Silkworm, *Bombyx mori*. *PLoS ONE* 2016, 11, e0147147.
145. Zhang, X.; Yuan, J.; Sun, Y.; Li, S.; Gao, Y.; Yu, Y.; Liu, C.; Wang, Q.; Lv, X.; Zhang, X.; et al. Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. *Nat. Commun.* 2019, 10, 356.
146. Gutekunst, J.; Andriantsoa, R.; Falckenhayn, C.; Hanna, K.; Stein, W.; Rasamy, J.; Lyko, F. Clonal genome evolution and rapid invasive spread of the marbled crayfish. *Nat. Ecol. Evol.* 2018, 2, 567–573.
147. Tang, B.; Zhang, D.; Li, H.; Jiang, S.; Zhang, H.; Xuan, F.; Ge, B.; Wang, Z.; Liu, Y.; Sha, Z.; et al. Chromosome-level genome assembly reveals the unique genome evolution of the swimming crab (*Portunus trituberculatus*). *Gigascience* 2020, 9, giz161.
148. Chebbi, M.A.; Becking, T.; Moumen, B.; Giraud, I.; Gilbert, C.; Peccoud, J.; Cordaux, R. The Genome of *Armadillidium vulgare* (Crustacea, Isopoda) Provides Insights into Sex Chromosome Evolution in the Context of Cytoplasmic Sex Determination. *Mol. Biol. Evol.* 2019, 36, 727–741.
149. Thomas, G.W.C.; Dohmen, E.; Hughes, D.S.T.; Murali, S.C.; Poelchau, M.; Glastad, K.; Anstead, C.A.; Ayoub, N.A.; Batterham, P.; Bellair, M.; et al. Gene content evolution in the arthropods. *Genome Biol.* 2020, 21, 15.
150. Pang, Y.; Mao, C.; Liu, S. Encoding activities of non-coding RNAs. *Theranostics* 2018, 8, 2496–2507.
151. Mori, K.; Weng, S.M.; Arzberger, T.; May, S.; Rentzsch, K.; Kremmer, E.; Schmid, B.; Kretzschmar, H.A.; Cruts, M.; Van Broeckhoven, C.; et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science* 2013, 339, 1335–1338.
152. Wang, X.Q.; Dostie, J. Reciprocal regulation of chromatin state and architecture by HOTAIRM1 contributes to temporal collinear HOXA gene activation. *Nucleic Acids Res.* 2017, 45, 1091–1104.
153. Jenkins, A.M.; Waterhouse, R.M.; Muskavitch, M.A. Long non-coding RNA discovery across the genus *Anopheles* reveals conserved secondary structures within and beyond the *Gambiae* complex. *BMC Genom.* 2015, 16, 337.
154. Legeai, F.; Derrien, T. Identification of long non-coding RNAs in insects genomes. *Curr. Opin. Insect Sci.* 2015, 7, 37–44.
155. Chang, Z.X.; Ajayi, O.E.; Guo, D.Y.; Wu, F. Genome-wide characterization and developmental expression profiling of long non-coding RNAs in *Sogatella furcifera*. *Insect Sci.* 2020, 27, 987–997.
156. Guan, R.; Li, H.; Zhang, H.; An, S. Comparative analysis of dsRNA-induced lncRNAs in three kinds of insect species. *Arch. Insect Biochem. Physiol.* 2020, 103, e21640.
157. Choudhary, C.; Sharma, S.; Meghwanshi, K.K.; Patel, S.; Mehta, P.; Shukla, N.; Do, D.N.; Rajpurohit, S.; Suravajhala, P.; Shukla, J.N. Long Non-Coding RNAs in Insects. *Animals* 2021, 11, 1118.
158. Ronshaugen, M.; Biemar, F.; Piel, J.; Levine, M.; Lai, E.C. The *Drosophila* microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes Dev.* 2005, 19, 2947–2952.
159. Garaulet, D.L.; Lai, E.C. Hox miRNA regulation within the *Drosophila* Bithorax complex: Patterning behavior. *Mech. Dev.* 2015, 138, 151–159.
160. Hui, J.H.; Marco, A.; Hunt, S.; Melling, J.; Griffiths-Jones, S.; Ronshaugen, M. Structure, evolution and function of the bi-directionally transcribed iab-4/iab-8 microRNA locus in arthropods. *Nucleic Acids Res.* 2013, 41, 3352–3361.
161. Negre, B.; Ruiz, A. HOM-C evolution in *Drosophila*: Is there a need for Hox gene clustering? *Trends Genet.* 2007, 23, 55–59.
162. Shippy, T.D.; Ronshaugen, M.; Cande, J.; He, J.; Beeman, R.W.; Levine, M.; Brown, S.J.; Denell, R.E. Analysis of the *Tribolium* homeotic complex: Insights into mechanisms constraining insect Hox clusters. *Dev. Genes Evol.* 2008, 218, 127–139.
163. Zeng, D.; Chen, X.; Peng, J.; Yang, C.; Peng, M.; Zhu, W.; Xie, D.; He, P.; Wei, P.; Lin, Y.; et al. Single-molecule long-read sequencing facilitates shrimp transcriptome research. *Sci. Rep.* 2018, 8, 16920.

164. Wan, H.; Jia, X.; Zou, P.; Zhang, Z.; Wang, Y. The Single-molecule long-read sequencing of *Scylla paramamosain*. *Sci. Rep.* 2019, 9, 12401.
165. Kim, H.S.; Kim, B.M.; Lee, B.Y.; Souissi, S.; Park, H.G.; Lee, J.S. Identification of Hox genes and rearrangements within the single homeobox (Hox) cluster (192.8 kb) of the cyclopoid copepod (*Paracyclopina nana*). *J. Exp. Zool. B Mol. Dev. Evol.* 2016, 326, 105–109.
166. Kim, D.H.; Lee, B.Y.; Kim, H.S.; Jeong, C.B.; Hwang, D.S.; Kim, I.C.; Lee, J.S. Identification and characterization of homeobox (Hox) genes and conservation of the single Hox cluster (324.6 kb) in the water flea *Daphnia magna*. *J. Exp. Zool. B Mol. Dev. Evol.* 2018, 330, 76–82.
167. Chipman, A.; Ferrier, D.; Brena, C.; Qu, J.; Hughes, D.S.T.; Schröder, R.; Torres-Oliva, M.; Znassi, N.; Jiang, H.; Almeida, F.; et al. The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. *PLoS Biol.* 2014, 12, e1002005.
168. Qu, Z.; Nong, W.; So, W.L.; Barton-Owen, T.; Li, Y.; Leung, T.C.N.; Li, C.; Baril, T.; Wong, A.Y.P.; Swale, T.; et al. Millipede genomes reveal unique adaptations during myriapod evolution. *PLoS Biol.* 2020, 18, e3000636.
169. Bakalenko, N.I.; Novikova, E.L.; Nesterenko, A.Y.; Kulakova, M.A. Hox gene expression during postlarval development of the polychaete *Alitta virens*. *Evodevo* 2013, 4, 13.
170. Novikova, E.L.; Bakalenko, N.I.; Nesterenko, A.Y.; Kulakova, M.A. Expression of Hox genes during regeneration of nereid polychaete *Alitta (Nereis) virens* (Annelida, Lophotrochozoa). *Evodevo* 2013, 4, 14.
171. Novikova, E.L.; Bakalenko, N.I.; Kulakova, M.A. Annelids win again: The first evidence of Hox antisense transcription in Spiralia. *bioRxiv* 2021.
172. Andreeva, T.F.; Cook, C.; Korchagina, N.M.; Akam, M.; Dondya, A.K. Cloning and analysis of structural organization of Hox genes in the polychaete *Nereis virens*. *Ontogenez* 2001, 32, 225–233. (In Russian)
173. Deschamps, J.; Duboule, D. Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes Dev.* 2017, 31, 1406–1416.
174. Imura, T.; Pourquié, O. Hox genes in time and space during vertebrate body formation. *Dev. Growth Differ.* 2007, 49, 265–275.
175. Hara, Y.; Yamaguchi, K.; Onimaru, K.; Kadota, M.; Koyanagi, M.; Keeley, S.D.; Tatsumi, K.; Tanaka, K.; Motone, F.; Kageyama, Y.; et al. Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nat. Ecol. Evol.* 2018, 2, 1761–1771.
176. Murray, S.C.; Mellor, J. Using both strands: The fundamental nature of antisense transcription. *Bioarchitecture* 2016, 6, 12–21.
177. Xu, Z.; Wei, W.; Gagneur, J.; Clauder-Münster, S.; Smolik, M.; Huber, W.; Steinmetz, L.M. Antisense expression increases gene expression variability and locus interdependency. *Mol. Syst. Biol.* 2011, 7, 468.
178. Fusco, G. Trunk segment numbers and sequential segmentation in myriapods. *Evol. Dev.* 2005, 7, 608–617.