Current Gene Therapy, 2018, *18*, 275-285

RESEARCH ARTICLE

EZH2 RIP-seq Identifies Tissue-specific Long Non-coding RNAs

Yan Wang[1,#], Yinping Xie[2,#], Lili Li[2,#], Yuan He[2], Di Zheng[3], Pengcheng Yu[2], Ling Yu[3], Lixu Tang[4], Yibin Wang[5,*] and Zhihua Wang[2,*]

[1]*Department of Cardiovascular Medicine, Beijing Hospital, National Center of Gerontology, Beijing 100730,China;* [2]*Department of Cardiology, Central Laboratory, Renmin Hospital, Wuhan University, Wuhan 430060, China;* [3]*Department of Orthopedics, Renmin Hospital, Wuhan University, Wuhan 430060, China;* [4]*Wushu College, Wuhan Sports University, Wuhan, Hubei 430079, China;* [5]*Departments of Anesthesiology, Division of Molecular Medicine, Physiology and Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA*

ARTICLE HISTORY

Received: March 25, 2018
Revised: May 24, 2018
Accepted: September 17, 2018

DOI:
10.2174/1566523218666181008125010

Abstract: ***Background*:** Polycomb Repressive Complex 2 (PRC2) catalyzes histone methylation at H3 Lys27, and plays crucial roles during development and diseases in numerous systems. Its catalytic subunit EZH2 represents a key nuclear target for long non-coding RNAs (lncRNAs) that emerging to be a novel class of epigenetic regulator and participate in diverse cellular processes. LncRNAs are characterized by high tissue-specificity; however, little is known about the tissue profile of the EZH2-interacting lncRNAs.

***Objective*:** Here we performed a global screening for EZH2-binding lncRNAs in tissues including brain, lung, heart, liver, kidney, intestine, spleen, testis, muscle and blood by combining RNA immuno-precipitation and RNA sequencing. We identified 1328 EZH2-binding lncRNAs, among which 470 were shared in at least two tissues while 858 were only detected in single tissue. An RNA motif with specific secondary structure was identified in a number of lncRNAs, albeit not in all EZH2-binding lncRNAs. The EZH2-binding lncRNAs fell into four categories including intergenic lncRNA, antisense lncRNA, intron-related lncRNA and promoter-related lncRNA, suggesting diverse regulations of both *cis* and *trans*-mechanisms. A promoter-related lncRNA Hnf1aos1 bound to EZH2 specifically in the liver, a feature same as its paired coding gene Hnf1a, further confirming the validity of our study. In addition to the well known EZH2-binding lncRNAs like Kcnq1ot1, Gas5, Meg3, Hotair and Malat1, majority of the lncRNAs were firstly reported to be associated with EZH2.

***Conclusion*:** Our findings provide a profiling view of the EZH2-interacting lncRNAs across different tissues, and suggest critical roles of lncRNAs during cell differentiation and maturation.

Keywords: Epigenetics, Long non-coding RNA, Tissue specificity, PRC2, EZH2, Histone methylation.

1. INTRODUCTION

System biology methods including Genome-Wide Association Studies (GWAS) and high-throughput RNA sequencing have been widely used to dissect the mechanisms underpinning human diseases [1-4]. Despite these successes, the majority of genetic architecture and gene expression profile of human complex diseases remains unclear [5-8]. A major challenge in the post-genome era is to mine novel disease risks from multi-level omics data using combined system biology methods, which may expand our knowledge of the causes of genetic disease [9-11].

Over 90% of mammalian genome is actively transcribed from DNA, but only 2% is destined to code proteins [12, 13]. The remnant belongs to diverse categories of non-coding RNAs [14-16]. Traditional non-coding RNA categories including rRNAs, tRNAs, miRNAs, snRNAs and snoRNAs are mainly transcribed by RNA polymerase I or III, while mRNAs are transcribed by polymerase II. Although it has long been recognized that Pol II-transcribed long non-coding RNAs (lncRNAs) exist, their critical functions in diverse cellular processes have not been explored until recently [17, 18].

Consistent with notion that expression of lncRNAs is highly tissue specific [19], emerging roles of lncRNAs in the development of certain organs have been reinforced [12, 20]. H19 is an important regulator of mammalian development and disease in that it inhibits cell proliferation [21]. Braveheart (Bvht), by modulating the core cardiovascular gene network, is necessary to maintain cardiac commitment [22]. Conversely, the lateral mesoderm-specific lncRNA Fendrr (fetal-lethal non-coding developmental regulatory RNA) controls mesodermal differentiation, as well as heart and body wall development [23]. Six3OS acts in trans to regulate retinal development by modulating Six3 activity [24]. Lin-

*Address correspondence to these authors at the Renmin Hospital of Wuhan University, 99 Zhangzhidong Rd., Wuchang District, Wuhan, 430060, China; Tel/Fax: +86-27-88041919; E-mail: zhihuawang@whu.edu.cn and Division of Molecular Medicine, Departments of Anesthesiology, Physiology and Medicine, David Geffen School of Medicine, University of California at Los Angeles, CA 90095, USA;
E-mail: yibinwang@mednet.ucla.edu

These authors contribute to this work equally.

1875-5631/18 $58.00+.00

cRNA1230 (linc1230) is both necessary and sufficient to repress neural commitment of mouse ES cells [25]. The smooth muscle and endothelial cell-enriched migration/differentiation-associated lncRNA (SENCR) is identified in human vascular smooth muscle and endothelial cells, being involved in their differentiation [26, 27]. These findings emphasize that lncRNAs are crucial for the cell fate determination during development.

Mechanistically, lncRNA functions as a signal, decoy, scaffold or guide due to its unlimited capability to bind to DNA, RNA or protein molecules [12]. Interestingly, majority of the lncRNAs locates inside the nucleus and interact with epigenetic modifiers, particularly, the polycomb repressive complex 2 (PRC2) [17, 22, 23, 28-34]. PRC2 catalyzes the tri-methylation at histone H3 lysine 27 (H3K27), and leads to chromatin remodeling to silence gene expression [35]. Though nearly all PRC2 subunits have potential to bind RNAs, the catalytic subunit EZH2 (enhancer of zeste homolog 2) has the highest affinity and is most frequently reported to be the molecular target of numerous lncRNAs [36]. A screening in embryonic stem cells using RNA Immune-Precipitation (RIP) method identifies over 9000 EZH2-binding RNAs [37], suggesting an indispensible feature of RNAs in EZH2-mediated gene programming. Nevertheless, how lncRNAs modify the function of EZH2 during development remains elusive.

Here, we performed an unbiased screening for EZH2-binding lncRNAs using RIP-seq in ten tissues including brain, lung, heart, liver, kidney, intestine, spleen, testis, muscle and blood. We identified both common EZH2-binding lncRNAs shared by diverse tissues and tissue-specific ones that potentially maintaining the differentiated cell status. Our study provides a comprehensive understanding of the molecular function of EZH2 and its related lncRNAs.

2. MATERIALS AND METHODS

2.1. Animal Approval

One male 2-month old C57/BL6 mouse was sacrificed by dislocation of infra-cervical spine. Tissues including brain, lung, heart, liver, kidney, spleen, intestine, skeletal muscle, testis and blood cells were quickly separated, washed in PBS, frozen in liquid nitrogen until use. All animal protocols were reviewed and approved by the Animal Care and Use Committee of Renmin Hospital at Wuhan University and conformed to the Guide for the Care and Use of Laboratory Animals, published by the US National Institutes of Health [38].

2.2. RNA Immune-precipitation

RNA immune-precipitation (RIP) was performed essentially as described [33, 39]. 200-300 mg of mouse tissues including brain, heart, lung, liver, kidney, spleen, intestine, testis, muscle and blood cells were homogenized in adequate volumes of polysome lysis buffer (10 mM HEPES-KOH (pH 7.0), 100 mM KCl, 5 mM $MgCl_2$, 25 mM EDTA, 0.5% IGEPAL, 2 mM dithiothreitol (DTT), 0.2 mg/mL Heparin, 50 U/mL RNase OUT (Life Technologies, NY, USA), 50 U/mL Superase IN (Ambion) and 1× complete protease inhibitor tablet (Roche)). The suspension was centrifuged at 14,000 *g* at 4 °C for 10 min to remove debris. Lysates containing 1 mg protein were incubated with 500 ng normal IgG (Cell Signaling Technologies, MA, USA; #2729, 1:200) or anti-EZH2 (Cell Signaling Technologies, MA, USA; #5246, 1:200) at 4 °C overnight on an inverse rotator. Protein A-sepharose beads (Life Technologies, 50 mL per tube) were first blocked in NT2 buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM $MgCl_2$ and 0.05% IGEPAL) supplemented with 5% BSA, 0.02% sodium azide and 0.02 mg/mL heparin at 4 °C for 1 h, and then added into the lysates followed by a 3-h incubation at 4 °C on an inverse rotator. The beads were subsequently washed five times in NT2 buffer. RNAs were released by incubating in proteinase K buffer (50 mM Tris (pH 8.0), 100 mM NaCl, 10 mM EDTA, 1% SDS and 1 U/mL proteinase K) for 30 min at 65 °C, and pelleting by adding an equal volume of isopropanol and centrifuging at 12,000*g* at 4 °C for 10 min. RNAs were washed once with 75% ethanol and stored at -80 °C until use.

2.3. Real-time PCR

Briefly, 1 mg RNA was reverse-transcribed into first-strand cDNA using the Superscript III first-strand synthesis kit (Life Technologies, NY, USA) with random primers. Real-time PCR was performed using the CFX96 Real-Time PCR Detection System (Bio-Rad, CA, USA) using the iQ SYBR Green Supermix (Bio-Rad). Values were normalized to IgG controls.

2.4. RNA Sequencing

Purified RIP RNAs were reverse transcribed into cDNA sequencing library using KAPA Stranded RNA-Seq Library Preparation Kit. The libraries were subjected to quality validation using the Agilent Bioanalyzer 2100, and sequenced using Illumina NextSeq 500 in DNA Link USA Inc.. The reads were mapped to mouse genome (mm10) using TopHat2 [33], and visualized on the UCSC browser (http://genome.ucsc.edu). LncRNAs were picked out according to NONCODE database [40]. Screening criteria was set as reads > 1.0 in at least one tissue; ratio of anti-EZH2 group relative to normal IgG group > 1.5.

2.5. In Silicon RNA Secondary Structure Prediction

RNA secondary structure was predicted by RNAfold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) based on Minimum Free Energy (MFE) and partition function.

2.6. Statistics

Comparisons in multiple groups were analyzed with one-way ANOVA. Data are presented as mean ± s.d.

3. RESULTS

3.1. Establishment of RIP Method in Different Tissues

RNA immune-precipitation (RIP) was performed using 200-300 mg tissues with anti-EZH2 antibody and normal IgG according to previous reports [33, 39]. Before RNA-seq, RT-PCR was used to detect known EZH2-binding lncRNAs to validate the success of RIP method. We measured the enrichment of three lncRNAs, *i.e.* cardiac hypertrophy associated epigenetics regulator (Chaer), HOX transcript antisense

RNA (Hotair) and H19. The results showed that Chaer was substantially enriched in heart, spleen and testis (>2 folds; Fig. **1A**), which was consistent with our previous finding that Chaer is a heart-specific lncRNA with detectable expression in spleen [33]. Whereas Hotair was substantially enriched in heart and blood (Fig. **1B**), H19 was detected in heart, muscle and blood (Fig. **1C**). These data confirm the success of current RIP method to enrich EZH2-binding lncRNAs.

3.2. RIP-seq Identifies EZH2-interacting lncRNAs

The pulled-down RNAs were then constructed to cDNA library using KAPA Stranded RNA-Seq Library Preparation Kit, and sequenced using Illumina NextSeq 500. The reads were mapped to mouse genome (mm10). LncRNAs were picked out according to NONCODE database [40]. After screening for reads no less than 1.0 and a ratio over 1.5 enrichment (anti-EZH2 group relative to IgG control), we obtained totally 1328 EZH2-binding lncRNAs in all ten tissues (Table **1**). Spleen is the tissue with the most EZH2-binding lncRNAs. Whereas 470 lncRNAs were shared in at least two tissues, 858 lncRNAs (64.61%) were tissue-specific EZH2-binding lncRNAs (Table **1**). Most lncRNAs fell into four categories based on their positions in the genome; *i.e.* intergenic lncRNA, antisense lncRNA, intron-related lncRNA and promoter-related lncRNA (Fig. **2**), suggesting an involvement of both *cis* and *trans* regulatory mechanisms. Among the identified lncRNAs, several were well-known EZH2-interacting lncRNAs including KCNQ1 overlapping transcript 1 (Kcnq1ot1) [28], growth arrest specific 5 (Gas5) [41], maternally expressed 3 (Meg3) [42] and Hotair [29] (Fig. **3A-D**), further confirming the veracity of the study. Gm12840 was a lncRNA detected in 5 tissues (Fig. **3E**). We indeed identified a common motif with paired two 4-nt loop secondary structure (Fig. **3F**) that was a typical feature responsible for the interaction with EZH2 [33, 43].

3.3. EZH2-interacting lncRNAs Shared in Different Tissues

The common lncRNAs identified in different tissues could serve as self-proof candidates. We identified 64 lncRNAs shared by at least three tissues (Table **2**), including the well-established lncRNAs, metastasis associated lung adenocarcinoma transcript 1 (Malat1) and Meg3 [30, 42, 44-47]. Malat1 has been reported to regulate diverse biological processes including development, differentiation and

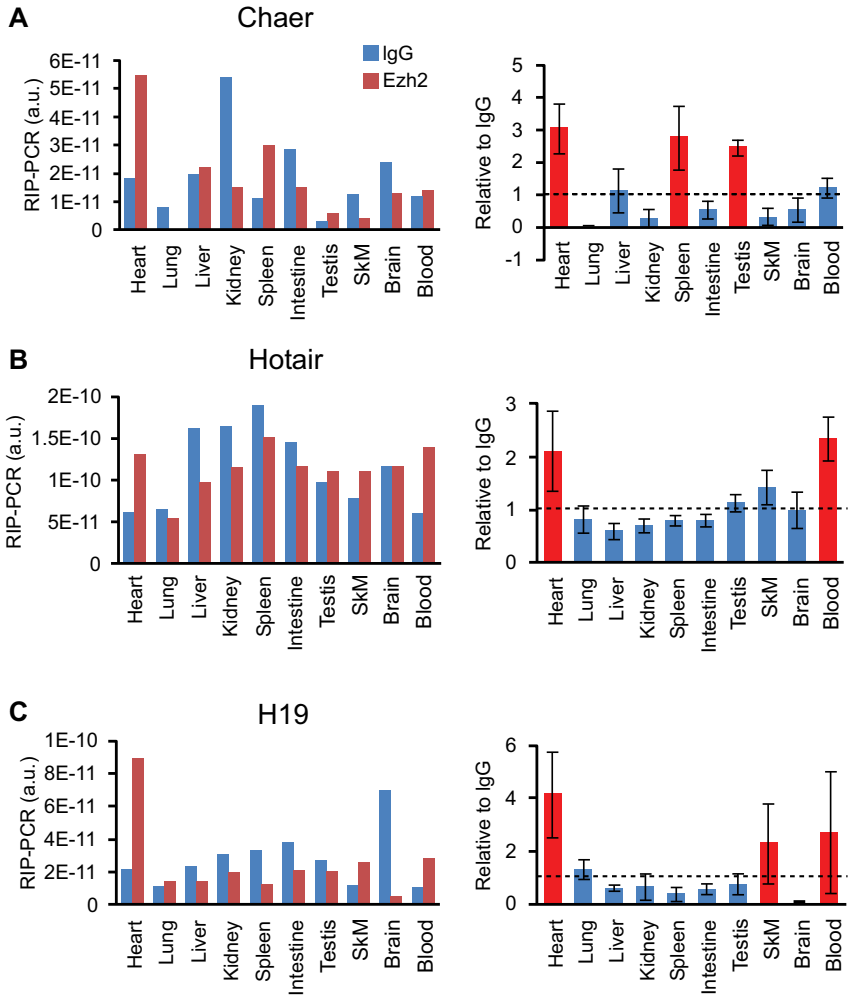

Fig. (1). Validation of the RNA immune-precipitation samples. (**A-C**) Real-time PCR was used to detect EZH2-binding lncRNAs including Chaer (**A**), Hotair (**B**) and H19 (**C**), with the arbitrary values in normal IgG and anti-EZH2 groups in left and ratio of anti-EZH2 to IgG in right. Data were mean ± SD. SKM: skeletal muscle.

Table 1. EZH2-binding lncRNAs in different tissues.

Tissues	EZH2-binding LncRNAs	With Tissue Specificity	Percentage (%)
Brain	87	38	43.68
Lung	41	6	14.63
Heart	52	23	44.23
Liver	76	36	47.37
Spleen	582	488	83.85
Kidney	62	29	46.77
Intestine	150	98	65.33
Testis	131	80	61.07
Muscle	112	53	47.32
Blood	35	7	20.00
Total	1328	858	64.61

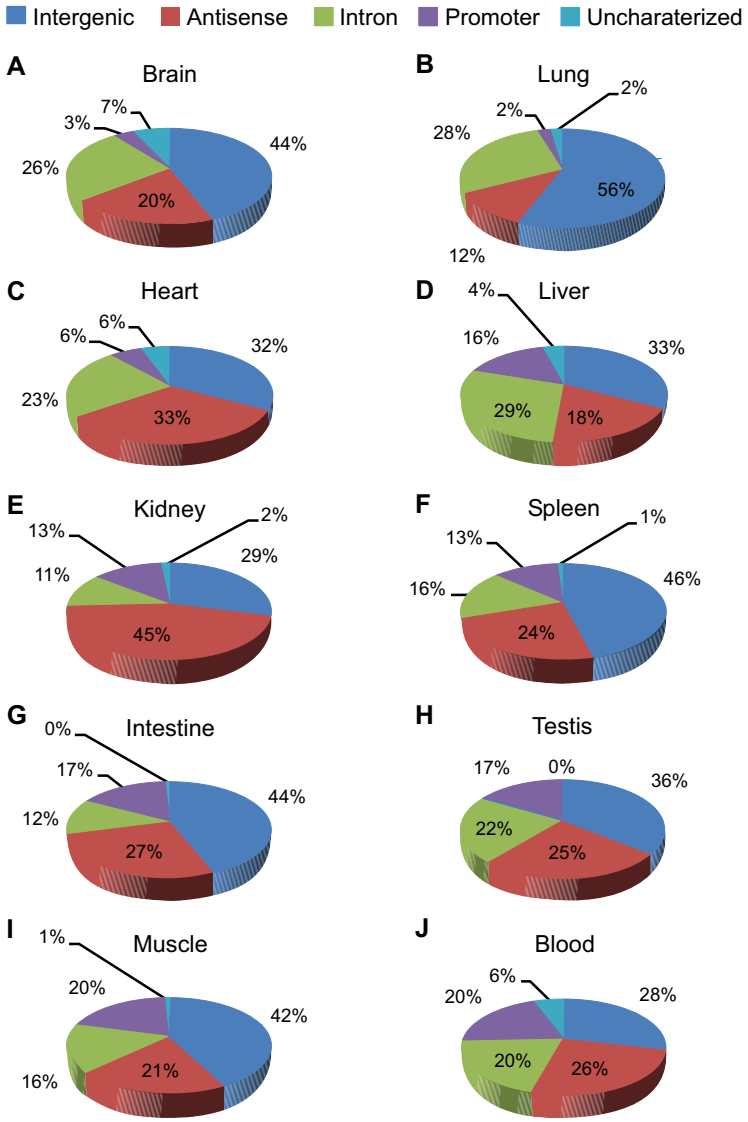

Fig. (2). Distribution of lncRNA categories in different tissues. (**A-J**) Pie chart analyses for EZH2-binding lncRNAs classified into intergenic, antisense, intron and promoter related lncRNAs in brain (**A**), lung (**B**), heart (**C**), liver (**D**), kidney (**E**), spleen (**F**), intestine (**G**), testis (**H**), muscle (**I**) and blood (**J**).

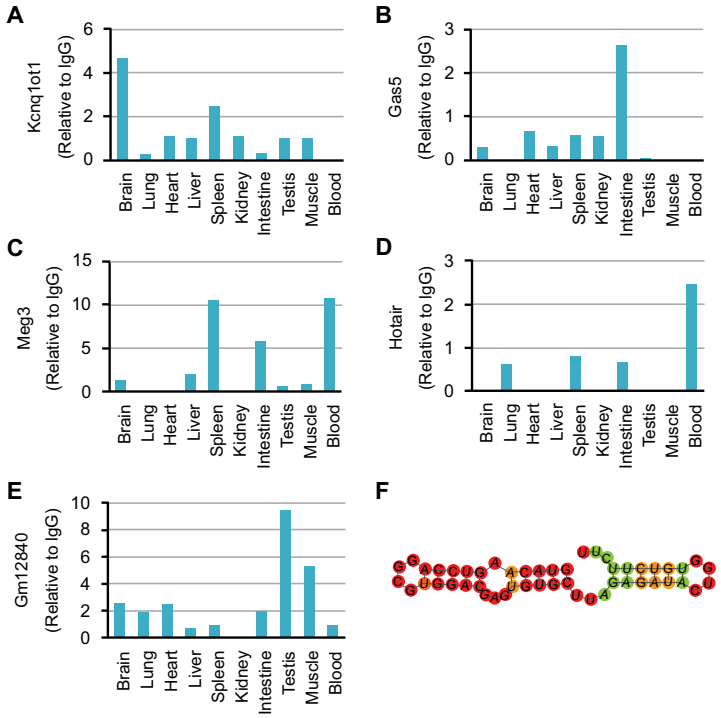

Fig. (3). Validated EZH2-binding lncRNAs. (**A-E**) Ratio of reads in anti-EZH2 group relative to IgG control for EZH2-binding lncRNAs, Kcnq1ot1 (**A**), Gas5 (**B**), Meg3 (**C**), Hotair (**D**) and Gm12840 (**E**) in different tissues. (**F**) Secondary RNA structure of a paired 4-nt loop motif responsible for the interaction with EZH2 in Gm12840.

Table 2. EZH2-binding lncRNAs shared in at least 3 tissues.

Gene Name	Detected Tissue									
	Bn	**Lg**	**Ht**	**Lv**	**Sp**	**Kd**	**In**	**Ts**	**Mu**	**Bd**
Gm37494	Y	-	-	-	Y	Y	Y	Y	Y	-
Gm37917	Y	-	Y	Y	Y	-	-	Y	-	-
Gm17131	Y	-	-	Y	Y	-	-	-	Y	Y
Gm13727	-	-	Y	Y	Y	-	-	Y	Y	-
Gm12840	-	-	Y	Y	Y	-	-	-	Y	Y
Snord13	-	-	-	Y	-	-	Y	Y	Y	-
Snora23	-	Y	-	Y	-	-	-	Y	Y	-
Snhg20	-	-	-	Y	-	-	Y	Y	Y	-
Gm27206	Y	-	Y	-	-	-	-	-	Y	Y
Gm25776	-	-	Y	Y	-	Y	-	-	Y	-
Gm25395	Y	Y	-	Y	-	-	-	-	Y	-
Gm25117	Y	-	-	Y	-	-	-	Y	-	Y
Gm24407	-	-	-	Y	-	Y	-	Y	Y	-
Gm23143	-	-	-	-	-	-	Y	Y	Y	Y

(Table 2) contd….

Gene Name	Detected Tissue									
	Bn	Lg	Ht	Lv	Sp	Kd	In	Ts	Mu	Bd
Gm22513	-	-	-	-	-	Y	Y	Y	Y	-
Gm20528	-	-	-	-	Y	Y	Y	Y	-	-
Uckl1os	Y	-	-	-	-	-	Y	Y	-	-
Snora57	-	Y	-	Y	-	-	-	-	Y	-
Snhg12	-	Y	-	-	-	-	Y	Y	-	-
Meg3	-	-	-	-	Y	-	Y	-	-	Y
Malat1	Y	-	-	Y	-	-	-	-	-	Y
Gm9864	-	-	-	-	Y	Y	-	-	Y	-
Gm38393	-	Y	Y	-	-	Y	-	-	-	-
Gm38271	-	Y	-	Y	Y	-	-	-	-	-
Gm38194	-	Y	-	-	Y	-	-	-	Y	-
Gm37954	-	Y	-	-	Y	-	-	-	Y	-
Gm37601	-	Y	-	Y	Y	-	-	-	-	-
Gm37515	Y	-	-	Y	Y	-	-	-	-	-
Gm37376	Y	-	Y	Y	-	-	-	-	-	-
Gm37349	-	Y	-	-	Y	-	Y	-	-	-
Gm29055	Y	-	-	-	Y	-	-	Y	-	-
Gm29044	-	Y	-	-	Y	-	Y	-	-	-
Gm28268	Y	-	-	-	Y	-	-	-	-	Y
Gm27350	-	-	-	-	Y	-	-	-	Y	Y
Gm26917	-	Y	-	-	Y	-	-	-	Y	-
Gm26905	Y	-	-	-	Y	-	-	-	-	Y
Gm26870	Y	-	Y	-	-	-	-	-	-	Y
Gm26397	-	-	Y	-	-	-	-	-	Y	Y
Gm25939	-	-	-	Y	-	-	-	Y	Y	-
Gm25835	-	-	-	-	-	Y	-	Y	Y	-
Gm25099	-	-	-	-	-	-	Y	Y	Y	-
Gm24574	-	-	Y	-	Y	-	-	-	-	Y
Gm24265	-	Y	-	-	-	Y	-	-	Y	-
Gm23442	-	-	-	-	-	-	Y	Y	Y	-
Gm22486	-	-	-	-	-	Y	-	Y	Y	-
Gm22442	Y	-	-	-	-	Y	Y	-	-	-
Gm22285	Y	-	-	-	Y	-	-	-	-	Y
Gm17132	Y	-	Y	-	Y	-	-	-	-	-
Gm16579	-	Y	-	-	Y	-	-	-	-	Y
Gm15662	-	-	-	-	-	-	Y	-	Y	Y

(Table 2) contd….

Gene Name	Detected Tissue									
	Bn	**Lg**	**Ht**	**Lv**	**Sp**	**Kd**	**In**	**Ts**	**Mu**	**Bd**
Gm15635	-	-	-	Y	-	Y	-	-	Y	-
Gm15280	Y	Y	-	-	Y	-	-	-	-	-
Gm14493	-	-	-	Y	-	Y	-	Y	-	-
Gm13722	-	Y	-	Y	Y	-	-	-	-	-
Gm10516	-	-	-	-	Y	Y	Y	-	-	-
Gm10125	-	Y	-	Y	Y	-	-	-	-	-
Gm10011	Y	-	-	-	-	Y	-	-	Y	-
E130102H24Rik	-	-	Y	Y	-	Y	-	-	-	-
D930048G16Rik	-	-	-	-	-	Y	-	-	Y	Y
5830428H23Rik	Y	-	Y	Y	-	-	-	-	-	-
3110053B16Rik	Y	-	-	-	Y	-	Y	-	-	-
1700034H15Rik	-	-	-	-	Y	-	Y	Y	-	-
1700020I14Rik	-	-	-	-	Y	-	-	-	Y	Y
1600010M07Rik	-	-	-	-	Y	-	Y	Y	-	-

Bn: brain; Lg: lung; Ht: heart; Lv: liver; Sp: spleen; Kd: kidney; In: intestine; Ts: testis; Mu: muscle; Bd: blood. Y: present.

diseases [47-52]. At the gene locus of Malat1, two antisense lncRNAs are also expressed: Gm37376 at 5' region and Gm20417 at 3' region (Fig. **4A**). Interestingly, both of them were identified as EZH2-binding lncRNAs, and showing similar tissue-specific pattern as Malat1 (Fig. **4B-D**), implicating the complexity of epigenetic regulation by lncRNAs.

3.4. Tissue-specific EZH2-interacting lncRNAs

Despite of their low conservation among species, the expression of LncRNAs have been shown to be highly tissue specific [19], implying a role of lncRNAs in the determination of cell fate and the maintenance of differentiated cell function. Among the EZH2-binding lncRNAs identified in this study, 64.61% were detected in only one tissue. An example is Hnf1aos1, which locates at the promoter region of hepatocyte nuclear factor 1-alpha (Hnf1a; Fig. **5A**). Hnf1a is a hepatocyte-specific transcription activator required for the expression of several liver genes [53-56]. Consistently, Hnf1aos1was only detected in the liver (Fig. **5B**). There are three isoforms of Hnf1aos1 due to alternative splicing; and within the third one, we identified an EZH2-recoganizing motif with the typical two 4-nt loop structure (Fig. **5C**).

4. DISCUSSION

Histone methylation-mediated epigenetic barrier is the fundamental basis to explain tissue differentiation, which can hardly be reversed [57, 58]. Albeit that lncRNA has long been involved in epigenetic regulations, their functions in cell fate determination remain to be elucidated. Our findings using an unbiased screening for EZH2-binding lncRNAs in ten tissues provide a comprehensive understanding of the tissue-specific non-coding regulators of PRC2.

The interaction of PRC2 with lncRNAs has been shown to be promiscuous; *i.e.* thousands of RNA targets, including both lncRNAs and mRNAs, are identified as PRC2-interacting RNAs [37, 59, 60]. Even so, specific structural features are described to be responsible for the high-affinity interaction between lncRNAs and PRC2 subunits. An 89-mer motif with two paired 4-nt loop "Crab-claw" structure was shown to bind EZH2-Embryonic Ectoderm Development (EED) dimer [43]. We previously reported a 66-mer motif with similar structure from the lncRNA cardiac hypertrophy associated epigenetics regulator (Chaer) was the molecular basis for its interaction with EZH2 [33]. In this study, the identified lncRNAs Gm12840 and Hnf1aos1 also possess similar structural motifs that may contribute to their interaction with EZH2. Nevertheless, majority of the lncRNAs do not show such obvious structural features. One possible reason is that lncRNAs might bind to EZH2 at different entity sites. Moreover, Kaneko *et al.* [30] reported that another component of PRC2 JARID2 assists with the interaction between PRC2 and lncRNAs. Ounzain *et al.* [61] found that lncRNA CARMEN interacts with both SUZ12 and EZH2 to regulate cardiomyocyte determination. In addition to PRC2, lncRNAs may simultaneously bind different factors to coordinate signaling transduction [62]. These evidence implies that other factors may modify the specificity and affinity of EZH2-lncRNA interaction.

More than half of the lncRNAs identified in our study are accompanying with a coding gene, either antisense or at the promoter region (Fig. **2**), suggesting an involvement of locally *cis*-regulation [30, 36]. It is still not clear whether such interaction keeps PRC2 away from the promoter or leads to spatial accessibility for histone modification. The fact that Hnf1aos1 and Hnf1a both exhibit liver-specific expression seemingly supports the latter. Han *et al.* [63] identified an lncRNA termed myheart (Mhrt), which locates at the 3' of the cardiac fetal gene beta-myosin heavy chain (Myh7). Under hypertrophic stress, Mhrt is upregualted and causes the induction of Myh7 through inhibiting the Brg1-Hdac-Parp

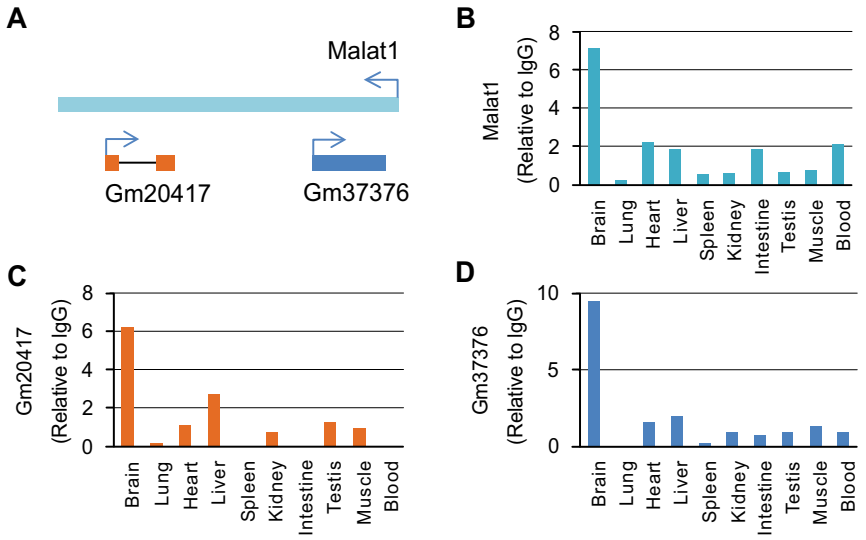

Fig. (4). EZH2 binds to Malat1 locus-derived lncRNAs. (**A**) Shematic of the Malat1 genomic structure together with two antisense lncRNAs, Gm20417 and Gm37376. (**B-D**) Ratio of reads in anti-EZH2 group relative to IgG control for Malat1 (**B**), Gm20417 (**C**) and Gm37376 (**D**) in different tissues.

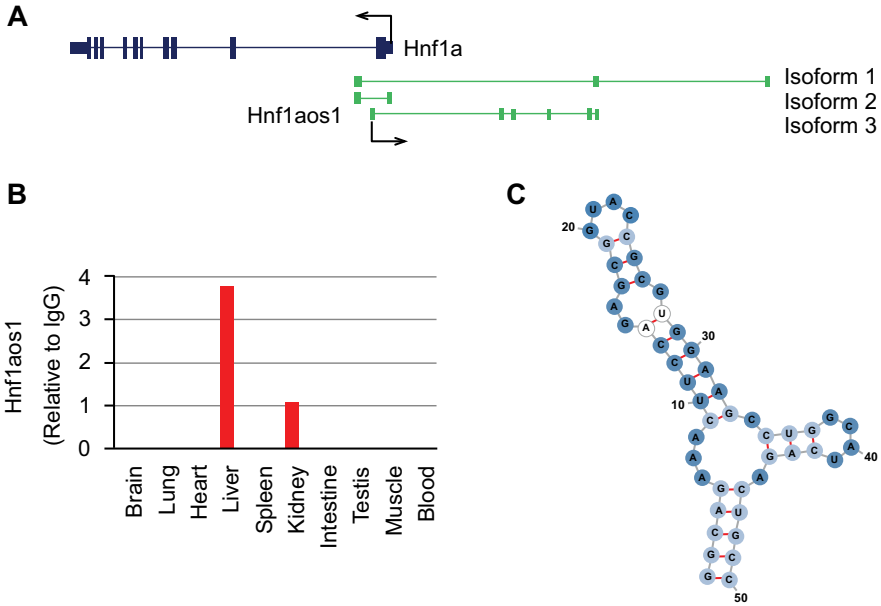

Fig. (5). An example showing a liver-specific EZH2-binding lncRNA locating near a liver-specific transcription activator. (**A**) Shematic of the genomic structure of lncRNA Hnf1aos1 and its neighbor gene Hnf1a. (**B**) Ratio of reads in anti-EZH2 group relative to IgG control for Hnf1aos1 in different tissues. (**C**) Predicted RNA structure of the motif identified in Hnf1aos1 responsible for its interaction with EZH2.

chromatin repressor complex. This evidence may provide a working model for EZH2-binding *cis*-regulatory lncRNAs. Importantly, the EZH2-binding lncRNAs turn to be abundant in the spleen compared with other tissues. This may not be a sequencing bias considering that the absolute reads in both anti-EZH2 group and the IgG control group resemble that observed in other tissues. The finding might imply a special requirement for lncRNA regulation in the spleen, the largest immune organ containing complicated cell types [64].

Though some studies suggest that lncRNAs might guide the locus-specific recruitment of epigenetic modifiers on genome [30, 65-72], other reports raised controversy mechanisms that lncRNAs might prevent the binding of PRC2 to specific gene promoters, and act as an activator of specific gene expression [73, 74]. Furthermore, the competition between different lncRNAs for the accessibility of available PRC2 apparatus should be carefully evaluated by researchers [74]. More efforts are required to investigate the molecular basis underpinning lncRNA-mediated gene regulation.

Epigenetic regulation is a highly dynamic process [75, 76]. In this study, all samples were isolated from one 2-month old adult male mouse. Due to experimental limitations *per se*, we did not perform experimental duplication on the RIP-seq assay. Nevertheless, EZH2-binding lncRNAs may vary with age, sex, circadian rhythm, and *etc*. Our study only reflects a snapshot of one epigenetic status. More details need to be accomplished to get a panoramic profiling of the lncRNAs surrounding PRC2 under different conditions in future.

With the development of novel gene delivery systems and gene editing toolkits, gene therapy has become a promising option to treat human diseases. One recent advancement is in Duchenne muscular dystrophy (DMD), which is caused by mutations in the X-linked dystrophin gene and is characterized by fatal degeneration of striated muscles. By using CRISPR-Cas9 and -Cpf1, the mutations could be corrected in human cells and animal disease model [77, 78]. In addition to inherited germline mutations, somatic mutations randomly occur at high rates and accumulate along with age [79, 80]. These mutations have been shown to contribute to the development of cancer and other lethal diseases [81]. Different tissues suffer from various environmental stresses that could lead to genetic alterations. The variance in genomic architecture define differential mutation hot spots across tissues where specific genes are expressed and differential DNA repair systems are implemented. Nevertheless, how to deliver genes to specific tissues without affecting others remains a technical barrier in this field. It would be favorable if the target gene displays high tissue specificity so that sequence-dependent treatment would not interfere with the normal function of tissues other than that hosting the target gene. To this end, microRNAs and siRNAs have been intensively explored as a therapeutic approach to silence causal genes of specific diseases. This field is further boosted by the prosperous nanoparticle mediated *in vivo* gene delivery [82]. So far, there is no reports about targeting lncRNAs in human diseases, albeit obvious advantages including high tissue specificity, scalable regulation on gene expression and low side effects. Considering that EZH2 is an important drug target for treatment of cancer and heart diseases, EZH2-binding lncRNAs can be leveraged to facilitate the medicine translation in clinic. Meanwhile, further investigations are required to clarify in detail the mechanisms how lncRNAs modify the epigenetic status and cell function.

Taken together, our findings reveal numerous tissue-specific EZH2-binding lncRNAs that display multiple interaction and regulation mechanisms. The data may help explain how an end-differentiated cell maintains its function and the genomic stability.

CONCLUSION

EZH2 RIP-seq identifies epigenetic lncRNAs with diverse genetic architectures. Whereas some lncRNAs are shared in multiple tissues, majority of the identified EZH2-binding lncRNAs show high tissue specificity, and may play an important role during organ development.

LIST OF ABBREVIATIONS

EZH2	=	Enhancer of Zeste Homolog 2
EED	=	Embryonic Ectoderm Development
JARID2	=	Jumonji, AT Rich Interactive Domain 2
PRC2	=	Polycomb Repressive Complex 2
RIP	=	RNA Immunoprecipitation
lncRNA	=	Long Non-coding RNA
H3K27me3	=	H3 Lys27 Tri-methylation
Chaer	=	Cardiac Hypertrophy Associated Epigenetics Regulator
Hotair	=	HOX Transcript Antisense RNA
Malat1	=	Metastasis Associated Lung Adenocarcinoma Transcript 1
Gas5	=	Growth Arrest Specific 5
Meg3	=	Maternally Expressed Gene 3
Kcnq1ot1	=	KCNQ1 Overlapping Transcript 1
Bvht	=	Braveheart
Mhrt	=	Myosin Heavy-Chain-Associated RNA Transcripts
Myh7	=	Myosin-7
Fendrr	=	FOXF1 Adjacent Non-Coding Developmental Regulatory RNA
Six3OS	=	Six3 Opposite Strand
Hnf1aos1	=	HNF1 Homeobox A, Opposite Strand 1
SENCR	=	Smooth Muscle and Endothelial Cell–Enriched Migration/Differentiation-Associated Long Non-coding RNA

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All animal protocols were reviewed and approved by the Animal Care and Use Committee of Renmin Hospital at Wuhan University.

HUMAN AND ANIMAL RIGHTS

No Human are used in the study. All animal research was conducted according to the Guide for the Care and Use of Laboratory Animals, published by the US National Institutes of Health.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

This study was supported by National Science and Technology Major Project (2017ZX10304402), Independent Research Project of Wuhan University (2042017kf0196), Excellent Young Investigator Project from National Natural Science Foundation of China (81722007), National Natural Science Foundation of China (NO:30900523, NO:81341013), Science and Technology Nova Plan of Beijing City (Z15111000030000, Z1511000003150100, 2015B068) and Science and Technology Nova Plan of Beijing Hospital (BJ-2016-035). We would like to thank Jon Choi, Gun Eui Lee and Anna Sheydina from DNA Link USA Inc., Christoph Rau, Ashely Cass and Xinshu Grace Xiao from UCLA for helping with the RNA-seq and data analysis.

REFERENCES

[1] Cheng L, Yang H, Zhao H, *et al*. MetSigDis: a manually curated resource for the metabolic signatures of diseases. Brief Bioinform 2017; doi: 10.1093/bib/bbx103. [Epub ahead of print].

[2] Hu Y, Zheng L, Cheng L, *et al*. GAB2 rs2373115 variant contributes to alzheimer's disease risk specifically in european population. J Neurol Sci 2017; 375: 18-22.

[3] Li Z, Hou P, Fan D, *et al*. The degradation of EZH2 mediated by lncRNA ANCR attenuated the invasion and metastasis of breast cancer. Cell Death Differ 2017; 24(1): 59-71.

[4] Peng J, Wang H, Lu J, *et al*. Identifying term relations cross different gene ontology categories. BMC Bioinformat 2017; 18(Suppl 16): 573.

[5] Cheng L, Jiang Y, Ju H, *et al*. InfAcrOnt: calculating cross-ontology term similarities using information flow by a random walk. BMC Genom 2018; 19(Suppl 1): 10.

[6] Liu G, Zhang F, Hu Y, *et al*. Genetic variants and multiple sclerosis risk gene slc9a9 expression in distinct human brain regions. Mol Neurobiol 2017; 54(9): 6820-6.

[7] Liu G, Zhang F, Jiang Y, *et al*. Integrating genome-wide association studies and gene expression data highlights dysregulated multiple sclerosis risk pathways. Mult Scler 2017; 23(2): 205-12.

[8] Liu G, Zhang Y, Wang L, *et al*. Alzheimer's disease rs11767557 variant regulates epha1 gene expression specifically in human whole blood. J Alzheimers Dis 2018; 61(3): 1077-88.

[9] Cheng L, Sun J, Xu W, *et al*. OAHG: an integrated resource for annotating human genes with multi-level ontologies. Sci Rep 2016; 6: 34820.

[10] Cheng L, Jiang Y, Wang Z, *et al*. DisSim: an online system for exploring significant similar diseases and exhibiting potential therapeutic drugs. Sci Rep 2016; 6: 30024.

[11] Pelletier R, Caron SO, Puymirat J. RNA based gene therapy for dominantly inherited diseases. Curr Gene Ther 2006; 6(1): 131-46.

[12] Devaux Y, Zangrando J, Schroen B, *et al*. Long noncoding RNAs in cardiac development and ageing. Nat Rev Cardiol 2015; 12(7): 415-25.

[13] Schonrock N, Jonkhout N, Mattick JS. Seq and you will find. Curr Gene Ther 2016; 16(3): 220-9.

[14] Guttman M, Garber M, Levin JZ, *et al*. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol 2010; 28(5): 503-10.

[15] Hu Y, Zhou M, Shi H, *et al*. Measuring disease similarity and predicting disease-related ncRNAs by a novel method. BMC Med Genomics 2017; 10(Suppl 5): 71.

[16] Cheng L, Hu Y, Sun J, *et al*. DincRNA: a comprehensive web-based bioinformatics toolkit for exploring disease associations and ncRNA function. Bioinformatics 2018; 34(11): 1953-6.

[17] Zhao J, Sun BK, Erwin JA, *et al*. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008; 322(5902): 750-6.

[18] Banerjee S, Smallwood A. A chromatin model of IGF2/H19 imprinting. Nat Genet 1995; 11(3): 237-8.

[19] Lee JH, Gao C, Peng G, *et al*. Analysis of transcriptome complexity through RNA sequencing in normal and failing murine hearts. Circ Res 2011; 109(12): 1332-41.

[20] Clark BS, Blackshaw S. Understanding the role of lncRNAs in nervous system development. Adv Exp Med Biol 2017; 1008: 253-82.

[21] Ratajczak MZ. Igf2-H19, an imprinted tandem gene, is an important regulator of embryonic development, a guardian of proliferation of adult pluripotent stem cells, a regulator of longevity, and a 'passkey' to cancerogenesis. Folia Histochem Cytobiol 2012; 50(2): 171-9.

[22] Klattenhoff CA, Scheuermann JC, Surface LE, *et al*. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell 2013; 152(3): 570-83.

[23] Grote P, Wittler L, Hendrix D, *et al*. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. Developmental cell 2013; 24(2): 206-14.

[24] Rapicavoli NA, Poth EM, Zhu H, Blackshaw S. The long noncoding RNA Six3OS acts in trans to regulate retinal development by modulating Six3 activity. Neural Dev 2011; 6: 32.

[25] Wang C, Li G, Wu Y, *et al*. LincRNA1230 inhibits the differentiation of mouse ES cells towards neural progenitors. Sci China Life Sci 2016; 59(5): 443-54.

[26] Boulberdaa M, Scott E, Ballantyne M, *et al*. A role for the long noncoding RNA SENCR in commitment and function of endothelial cells. Mol Ther 2016; 24(5): 978-90.

[27] Bell RD, Long X, Lin M, *et al*. Identification and initial functional characterization of a human vascular cell-enriched long noncoding RNA. Arterioscler Thromb Vasc Biol 2014; 34(6): 1249-59.

[28] Redrup L, Branco MR, Perdeaux ER, *et al*. The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. Development 2009; 136(4): 525-30.

[29] Tsai MC, Manor O, Wan Y, *et al*. Long noncoding RNA as modular scaffold of histone modification complexes. Science 2010; 329(5992): 689-93.

[30] Kaneko S, Bonasio R, Saldana-Meyer R, *et al*. Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. Mol Cell 2014; 53(2): 290-300.

[31] Mondal T, Subhash S, Vaid R, *et al*. MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. Nat Commun 2015; 6: 7743.

[32] Li X, Lin Y, Yang X, *et al*. Long noncoding RNA H19 regulates EZH2 expression by interacting with miR-630 and promotes cell invasion in nasopharyngeal carcinoma. Biochem Biophys Res Commun 2016; 473(4): 913-9.

[33] Wang Z, Zhang XJ, Ji YX, *et al*. The long noncoding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy. Nat Med 2016; 22(10): 1131-9.

[34] Liu Y, Peng B, Wu S, Xu N. Epigenetic regulation of regulatory T cells in kidney disease and transplantation. Curr Gene Ther 2017; 17(6): 461-8.

[35] Plath K, Fang J, Mlynarczyk-Evans SK, *et al*. Role of histone H3 lysine 27 methylation in X inactivation. Science 2003; 300(5616): 131-5.

[36] Cifuentes-Rojas C, Hernandez AJ, Sarma K, Lee JT. Regulatory interactions between RNA and polycomb repressive complex 2. Mol Cell 2014; 55(2): 171-85.

[37] Zhao J, Ohsumi TK, Kung JT, *et al*. Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell 2010; 40(6): 939-53.

[38] National Research Council. Guide for the care and use of laboratory animals 8th edn. the national academies Press 2011.

[39] Baroni TE, Chittur SV, George AD, Tenenbaum SA. Advances in RIP-chip analysis : RNA-binding protein immunoprecipitation-microarray profiling. Methods Mol Biol 2008; 419: 93-108.

[40] Zhao Y, Li H, Fang S, *et al*. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic Acids Res 2016; 44(D1): D203-8.

[41] Sun D, Yu Z, Fang X, *et al*. LncRNA GAS5 inhibits microglial M2 polarization and exacerbates demyelination. EMBO Rep 2017; 18(10): 1801-16.

[42] Terashima M, Tange S, Ishimura A, Suzuki T. MEG3 long noncoding rna contributes to the epigenetic regulation of epithelial-mesenchymal transition in lung cancer cell lines. J Biol Chem 2017; 292(1): 82-99.

[43] Wu L, Murat P, Matak-Vinkovic D, *et al*. Binding interactions between long noncoding RNA HOTAIR and PRC2 proteins. Biochemistry 2013; 52(52): 9519-27.

[44] Sanli I, Lalevee S, Cammisa M, *et al*. Meg3 Non-coding RNA expression controls imprinting by preventing transcriptional upregulation in cis. Cell Rep 2018; 23(2): 337-48.

[45] Iyer S, Modali SD, Agarwal SK. Long noncoding RNA MEG3 is an epigenetic determinant of oncogenic signaling in functional pancreatic neuroendocrine tumor cells. Mol Cell Biol 2017; 37(22): e00278-17.

[46] Das PP, Hendrix DA, Apostolou E, *et al*. PRC2 Is required to maintain expression of the maternal Gtl2-Rian-mirg locus by preventing De Novo DNA methylation in mouse embryonic stem cells. Cell Rep 2015; 12(9): 1456-70.

[47] Shi X, Sun M, Liu H, *et al*. Long non-coding RNAs: a new frontier in the study of human diseases. Cancer Lett 2013; 339(2): 159-66.

[48] Biswas S, Thomas AA, Chen S, *et al*. MALAT1: An epigenetic regulator of inflammation in diabetic retinopathy. Sci Rep 2018; 8(1): 6526.

[49] Dey BK, Mueller AC, Dutta A. Long non-coding RNAs as emerging regulators of differentiation, development, and disease. Transcription 2014; 5(4): e944014.

[50] Kim SH, Kim SH, Yang WI, *et al*. Association of the long non-coding RNA MALAT1 with the polycomb repressive complex pathway in T and NK cell lymphoma. Oncotarget 2017; 8(19): 31305-17.

[51] Qi Y, Ooi HS, Wu J, *et al*. MALAT1 long ncRNA promotes gastric cancer metastasis by suppressing PCDH10. Oncotarget 2016; 7(11): 12693-703.

[52] Wang D, Ding L, Wang L, *et al*. LncRNA MALAT1 enhances oncogenic activities of EZH2 in castration-resistant prostate cancer. Oncotarget 2015; 6(38): 41045-55.

[53] Ohtsuka H, Abe T, Onogawa T, *et al*. Farnesoid X receptor, hepatocyte nuclear factors 1alpha and 3beta are essential for

transcriptional activation of the liver-specific organic anion transporter-2 gene. J Gastroenterol 2006; 41(4): 369-77.

[54] Huang P, He Z, Ji S, *et al*. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature 2011; 475(7356): 386-9.

[55] Lee YH, Magnuson MA, Muppala V, Chen SS. Liver-specific reactivation of the inactivated Hnf-1alpha gene: elimination of liver dysfunction to establish a mouse MODY3 model. Mol Cell Biol 2003; 23(3): 923-32.

[56] Yamagata K, Oda N, Kaisaki PJ, *et al*. Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). Nature 1996; 384(6608): 455-8.

[57] Becker JS, Nicetto D, Zaret KS. H3K9me3-dependent heterochromatin: barrier to cell fate changes. Trends Genet 2016; 32(1): 29-41.

[58] Xie B, Zhang H, Wei R, *et al*. Histone H3 lysine 27 trimethylation acts as an epigenetic barrier in porcine nuclear reprogramming. Reproduction 2016; 151(1): 9-16.

[59] Wang X, Goodrich KJ, Gooding AR, *et al*. Targeting of polycomb repressive complex 2 to RNA by short repeats of consecutive guanines. Mol Cell 2017; 65(6): 1056-67 e5.

[60] Brockdorff N. Noncoding RNA and Polycomb recruitment. RNA 2013; 19(4): 429-42.

[61] Ounzain S, Micheletti R, Arnan C, *et al*. CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. J Mol Cell Cardiol 2015; 89(Pt A): 98-112.

[62] McHugh CA, Chen CK, *et al*. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 2015; 521(7551): 232-6.

[63] Han P, Li W, Lin CH, *et al*. A long noncoding RNA protects the heart from pathological hypertrophy. Nature 2014; 514(7520): 102-6.

[64] Jiang Q, Jin S, Jiang Y, *et al*. Alzheimer's disease variants with the genome-wide significance are significantly enriched in immune pathways and active in immune cells. Mol Neurobiol 2017; 54(1): 594-600.

[65] Liu GY, Zhao GN, Chen XF, *et al*. The long noncoding RNA Gm15055 represses Hoxa gene expression by recruiting PRC2 to the gene cluster. Nucleic Acids Res 2016; 44(6): 2613-27.

[66] Ma HW, Xie M, Sun M, *et al*. The pseudogene derived long noncoding RNA DUXAP8 promotes gastric cancer cell proliferation and migration *via* epigenetically silencing PLEKHO1 expression. Oncotarget 2017; 8(32): 52211-24.

[67] Marin-Bejar O, Marchese FP, Athie A, *et al*. Pint lincRNA connects the p53 pathway with epigenetic silencing by the Polycomb repressive complex 2. Genome Biol 2013; 14(9): R104.

[68] Negishi M, Wongpalee SP, Sarkar S, *et al*. A new lncRNA, APTR, associates with and represses the CDKN1A/p21 promoter by recruiting polycomb proteins. PLoS One 2014; 9(4): e95216.

[69] Nie FQ, Sun M, Yang JS, *et al*. Long noncoding RNA ANRIL promotes non-small cell lung cancer cell proliferation and inhibits apoptosis by silencing KLF2 and P21 expression. Mol Cancer Ther 2015; 14(1): 268-77.

[70] Ohhata T, Matsumoto M, Leeb M, *et al*. Histone H3 Lysine 36 trimethylation is established over the Xist promoter by antisense tsix transcription and contributes to repressing xist expression. Mol Cell Biol 2015; 35(22): 3909-20.

[71] O'Leary VB, Ovsepian SV, Carrascosa LG, *et al*. PARTICLE, a Triplex-Forming Long ncRNA, Regulates Locus-Specific Methylation in Response to Low-Dose Irradiation. Cell Rep 2015; 11(3): 474-85.

[72] Meredith EK, Balas MM, Sindy K, Haislop K, *et al*. An RNA matchmaker protein regulates the activity of the long noncoding RNA HOTAIR. RNA 2016; 22(7): 995-1010.

[73] Kim DH, Xi Y, Sung S. Modular function of long noncoding RNA, COLDAIR, in the vernalization response. PLoS Genet 2017; 13(7): e1006939.

[74] Elangovan VR, Camp SM, Kelly GT, *et al*. Endotoxin- and mechanical stress-induced epigenetic changes in the regulation of the nicotinamide phosphoribosyltransferase promoter. Pulmonary circulation 2016; 6(4): 539-44.

[75] Zhang B, Zheng H, Huang B, *et al*. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature 2016; 537(7621): 553-7.

[76] Kazakevych J, Sayols S, Messner B, *et al*. Dynamic changes in chromatin states during specification and differentiation of adult intestinal stem cells. Nucleic Acids Res 2017; 45(10): 5770-84.

[77] Zhang Y, Long C, Li H, *et al*. CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice. Sci Adv 2017; 3(4): e1602814.

[78] Long C, McAnally JR, Shelton JM, *et al*.Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science 2014; 345(6201): 1184-8.

[79] Xie M, Lu C, Wang J, *et al*. Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nature medicine 2014; 20(12): 1472-8.

[80] Kanchi KL, Johnson KJ, Lu C, *et al*. Integrated analysis of germline and somatic variants in ovarian cancer. Nat communicat 2014; 5: 3156.

[81] Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature 2009; 458(7239): 719-24.

[82] Min S, Jin Y, Hou CY, *et al*. Bacterial tRNase-based gene therapy with Poly(beta-Amino Ester) nanoparticles for suppressing melanoma tumor growth and relapse. Adv Healthc Mater 2018; 7(16): e1800052.

[83] Titze-de-Almeida R, Titze-de-Almeida SS. miR-7 replacement therapy in Parkinson's disease. Curr Gene Ther 2018; 18(3): 143-53.