Acta Biochim Biophys Sin 2012, 44: 40–47 | © The Author 2011. Published by ABBS Editorial Office in association with Oxford University Press on behalf of the Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. DOI: 10.1093/abbs/gmr112.

ABBS

Review

Genome-wide expression of non-coding RNA and global chromatin modification

Rukui Zhang[1,2], Lan Zhang[1,2], and Wenqiang Yu[1,2]*

[1]Key Laboratory of Ministry of Education, Department of Molecular Biology, Fudan University, Shanghai 200032, China
[2]EpiRNA Lab, Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China
*Correspondence address. Tel: +86-21-54237978; Fax: +86-21-54237339; E-mail: yuwqlizy@gmail.com

Traditionally, we know that genomic DNA will produce transcripts named messenger RNA and then translate into protein following the instruction of genetic central dogma, and RNA works here as a pass-by messenger. Now increasing evidence shows that RNA is a key regulator as well as a message transmitter. It is discovered by next-generation sequencing techniques that most genomic DNA are generally transcribed to non-coding RNA, highly beyond the percentage of coding mRNA. These non-coding RNAs (ncRNAs), belonging to several groups, have critical roles in many cellular processes, expanding our understanding of the RNA world. We review here the different categories of ncRNA according to genome location and how ncRNAs guide and recruit chromatin modification complex to specific loci of genome to modulate gene expression by affecting chromatin state.

Keywords non-coding RNA; chromatin modification complex; gene expression

Received: September 7, 2011 Accepted: November 10, 2011

Introduction

The complexity of the genes and the gene products in eukaryote is highly beyond our imagination. This is reflected not only in gigantic genome but also in its complicated transcription and regulation network. Intrinsically, the chromosomal DNA lives in different chamber-like regions, where the nucleosomes are tightly or less condensed, which is relevant to the chromatin accessibility for transcript factors [1]. Additionally, chromatin DNA partitioned to chromosome segregation is located in a specific region in nucleus with the help of lamin [2] and plenty of non-histone protein, and these lamina-associated spatial organizations likely represent large-scale heterochromatin formation with low transcription levels [3].

RNA transcription levels are largely monitored by chromatin state. Eukaryote genome is packed into a nucleosome that is composed of 146 bp DNA and core histones (H3, H4, H2A, and H2B). More importantly, histones are subject to an extensive range of post-translational modifications including methylation, acetylation, ubiquitination, phosphorylation, and polyADP ribosylation [4]. The genome DNA can be methylated [5] prevalently and hydroxymethylated [6] though in low level. Although DNA methylation affects the structure and dynamics of nucleosomes, usually resulting in a more compact and rigid structure; 5hmC are highly conjoined with promoters in embryonic stem cells, which usually present in poised genes with bivalent signatures marked by histone 3 lysine 27 trimethylation (H3K27me3) and histone 3 lysine 4 trimethylation (H3K4me3). Different histone modifications imply particular epigenetic functions. For example, H3K4me3 reside in promoter, H3K4me2 locate in promoters and enhancer regions, H3K4me1 sit on the enhancer, lysine 9 acetylation (H3K9ac), and H3K27ac correlate with active regulatory regions, H3K36me3 and H4K20me1 meet with transcribed regions and H3K27me3 which is associated with Polycomb-repressed regions [7,8]. Furthermore, combinations of two or more marks provide a more powerful strategy to identify new transcriptional units. Employing ‘K4–K36 domain’ discovered >1000 long non-coding RNAs (ncRNAs) which are conserved in mammals [9]. Considering the crucial roles of chromatin modification in gene expression during development and disease, we would like to ask: what factors are responsible for the initiation of these DNA and histone modification? How does the cell know in what time and which region of their genome should be modified? In this review, we propose that diversified ncRNAs transcribed from different genome regions such as enhancer, promoter, intron, and intergenic region guide and recruit chromatin modification complex to specific loci to affect chromatin state contributing to regulate gene expression in response to the environment alteration.

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024

Genome-wide Transcriptional ncRNA Bridging DNA and Chromatin Modification Complex

Genome DNA serves as carrier of static genetic information, which is interpreted by transcription and translation to manipulate phenotype. RNA polymerase II (Pol II), the well-known RNA polymerase consisting of 12 subunits [10], cooperates with a large number of transcription factors to elegantly initiate the transcription, and Pol II catalyzes DNA to produce heterogeneous RNA including ncRNA such as small nucleolar RNA (snoRNA), ribosomal RNA (rRNA) as well as traditional functional mRNA molecule that is translated into protein [9,11–13]. Specifically, promoter in eukaryote genome comprises enormous regulation elements to recruit various functional factors including Pol II and chromatin remolding complexes to change its configuration and adjust gene transcription. For a eukaryotic Pol II molecule, the 'correct' initiation transcription from a defined promoter leads to the production of protein-coding messenger RNA (mRNA), microRNA (miRNA) [14], endogenous small interfering RNA (siRNA) [15], piwi-interacting RNA (piRNA) [16], and large intergentic ncRNA (lincRNA) [9]. However, recently high-throughput sequencing shows that Pol II also initiates transcription 'non-specifically' from undefined or incorrect sites to generate promoter-associated short RNAs (PASR) and termini-associated short RNAs (TASR) and other uncharacterized RNAs with biological significance [17,18]. Moreover, genomic transcription is not always limited to active genes because any engaged Pol II may produce transcripts ranging in plenitude, which was confirmed by genome-wide mapping of Pol II uncovering that 10%–30% of the inactive genes are bound by Pol II in their promoter regions [19]. Even some short RNAs [20,21] are generated by paused Pol II when it is poising for but not embarking on functional transcription. Likewise, how Pol II is recruited to those promoters and whether Pol II always inhabits along DNA to scan proper promoters to initiate transcription or Pol II loiters or stably stalls at the promoter as prepared waiting for onset of stimuli are still undecided issues.

Recent studies by nuclear run-on sequencing [22] and NET-seq [23] and other powerful next-generation sequencing techniques [24] prove that most genomic DNA produces transcripts. In human genome, in addition to annotated transcripts from functional gene region which accounts for 1%–2% genome, there are plenty of unannotated transcripts [25], 80% of which stretch 93% genome [24] with relatively low-level copy number from particular loci. The observation that ncRNA transcripts extend from 5′ untranslated region (5′UTR) of elongation factor 1 alpha promoter (EF1a) [26] to gene body in low copy indicates that the promoter region is transcriptionally active. It seems that low-level transcripts are essential for gene silencing [27] and heterochromatin formation [20].

Genomic DNA can be separated into several functional units including promoter region, gene body, gene terminal region, and other regulated elements (**Fig. 1**), most of which transcribe ncRNA or become the target of ncRNA to modify chromatin structure followed by changes in gene expression. Unlike double-strand DNA, single-stand RNA molecule could easily fold into secondary structure like stem loop which could be recognized by coordinate proteins, and on the other hand the primary RNA sequence complements to chromosomal DNA via base pairing for targeting distinguishing loci. All these features render RNA as an excellent candidate to bridge DNA and associated modifying factors that remold chromatin, change chromatin accessibility, and manage gene expression. Considering that many ncRNAs are tightly associated with chromatin-modifying complex [28], the relationship between transcription, ncRNA, and chromatin modification is worthy of further exploration.

ncRNA Anchored to the Promoter Region and Chromatin Modification

Conceivably, promoter, not only with enormous transcriptional factors binding on it, but also being the target of ncRNAs which are discovered in increasing number, is a pivotal element to control gene expression. As Pol II initiates transcription from promoter, there leaves more opportunities to generate promoter-associated ncRNA, which possesses capacity to regulate gene expression in various epigenetic mechanisms. Firstly, the transcription process itself takes part in the arrangement of chromatin state as well as production of the transcripts, which can be verified by the fact that promoters of ncRNAs are more conserved than that of the protein-coding gene [29]. In this case, transcription allows promoter to obtain an open chromatin architecture that in turn benefits gene activation. Secondly, ncRNA allosterically regulates the activity of chromatin-associated factors [30]. Thirdly, the transcribing RNA serves as an adaptor to call for various factors to the chromatin complex when RNA molecule physically interacts with chromatin around the promoter region. Finally, the transcribed RNA may recognize its complementary DNA sequence so as to bring RNA-binding protein to the particular chromatin region [31].

Till now, most epigenetic events and gene regulation have focused on the promoter region including CpG island and CpG shore. Many genes employ multiple promoters to trigger transcription, along with a series of ncRNAs to control promoter activity. For instance, in human quiescent cells, the dihydrofolate reductase's upstream minorpromoter

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024

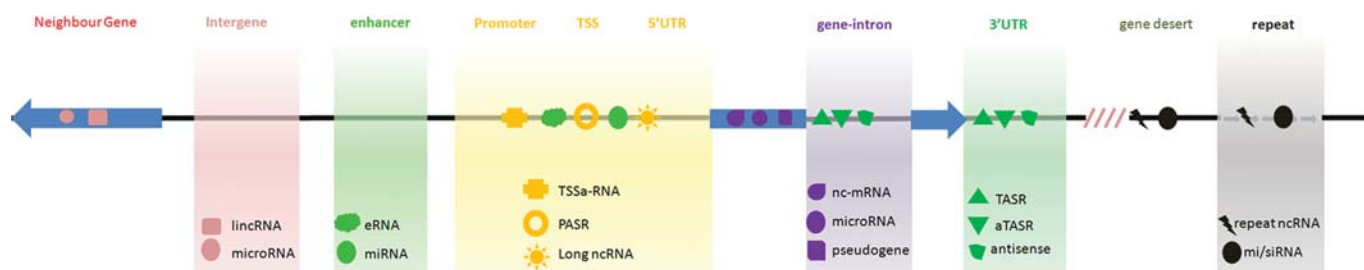

Figure 1 Genome-wide transcriptional landscape and global chromatin modification The top part shows functional units based on location around gene; the bottom part shows various types of ncRNA transcribed from the upper corresponding units; the shapes lying in DNA indicate targets of ncRNAs to modify chromatin structure. nc-mRNA: homologous to mRNA but do not code proteins; TSSa-RNA: TSS-associated RNA.

derives ncRNA to repress the major promoter-derived transcription that contributes to cell cycle regulation [27]. Imprinting genes usually have their antisense expression which may partially explain allele-specific expression. Taking insulin-like growth factor 2 (IGF2) as an example, it transcribes from four different promoters, designated P1–P4. While transcripts from P1 are derived from both parental alleles, transcripts from P2–P4 are from paternal allele. IGF2AS, one kind of ncRNA, is also a paternally expressed transcript produced within the IGF2 locus in a reverse orientation [32,33]. Does IGF2 antisense regulate the IGF2 expression paternally? Interestingly, IGF2 receptor gene that is maternally expressed is managed by an intronic imprint control element that contains the promoter region of its antisense namely Air that is paternally expressed to silence IGF2R by inducing methylation in the promoter region [34] in correlation with H3K9 methylation due to G9a recruitment by Air [35]. In X chromosome inactivation, different types of ncRNAs and chromatin modification complexes create a classical model to link ncRNA and chromatin state [36,37], which emerges as the universal pattern in an organism [28]. Even more ncRNAs probably lead to DNA methylation to sustain changed chromatin configuration during the cell division [38–42], which is similar to the well-established model of siRNA-directed DNA methylation in plants [43].

Different from siRNA-directed mRNA degradation, siRNA targeted in promoter region adjusts gene expression at transcriptional level by changing chromatin accessibility. The exogenous siRNA-directed gene silencing has been demonstrated in plants, yeasts, and mammalians through different mechanism to tune down gene transcription. What about endogenous siRNA in regulating transcription? One of the major recent scientific progresses is the discovery of a variety of ncRNAs surrounding the extensive defined promoter region of known genes transcribed in sense and antisense orientation [17,21,22,44,45], some of which may further be chopped into endogenous siRNA. An argonaute-containing complex RITS (RNA-induced initiation of transcriptional gene silencing) [46] is extremely tight to heterochromatin formation which holds imperative role in maintaining genomic stability. We anticipate that siRNA may recognize the promoter-associated RNAs which load to RITS complex to direct epigenetic modification on the corresponding targeted promoters. Many chromatin modification factors like DNA methyltransferases (DNMTs) and histone methyltransferases [38] are then recruited to the targeted elements. Chromatin immunoprecipitation (ChIP) assays show that H3K27me3 combining with DNMT3A and siRNA targeted to EF1promoter build a stable epigenetic modifications complex at the EF1 promoter.

SiRNA targeting to the promoter region does not always lead to gene silencing. Several distinct siRNAs designed to Cadherin-1 promoter can elevate the gene expression, although the mechanism is still controversial. The next-generation sequence technology also identifies a large number of transcripts surrounding mRNA [24,25]. Later, Yue *et al*. [47] verified that these extra ncRNA either inhibit or active gene expression. Addition of antigene RNA PR13515 that locates in region beyond 3′UTR and antigene RNA PR-11 that generate from the promoter region increase the amounts of progesterone receptor message transcripts (mRNA), respectively. miRNA is noted as one of the well-defined small ncRNA, and previous studies have shown that miRNAs regulate gene expression by translational suppression or sequence-specific degradation of complementary mRNA in post-transcriptional level. Nonetheless, recent research has unveiled that miR373 complement to E-cadherin promoter leading to Pol II enrichment at transcription start site (TSS) [48] and here miR373 emerged as a positive regulator in contrast to traditionally down-regulation gene expression by miRNA. Another example is 5′-RACE confirmed by the promoter-associated ncRNA derived from upstream of c-myc gene [49]. Essentially, those promoter-associated ncRNAs correlated positively with c-myc transcripts but ~50 times lower expression than c-myc mRNA. Using siRNA targeted to c-myc promoter-associated ncRNA which is seen in nuclear run-on assays can significantly reduce c-myc gene expression in cells.

In comparison with small ncRNAs, long ncRNAs are emerging as a new type of ncRNAs. Transcriptome profile

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024

by H3K4me3 and H3K36me3 enrichment in tiling array identified 216 ncRNAs that located in the regions 4–8 kb upstream of the TSS. One of them named PANDA interacts with the transcription factor NF-YA to mediate cell cycle arrest and apoptosis [50]. Surprisingly, the promoter region of ncRNA is even more conserved than that of the protein-coding mRNA in mammalian [29] which further strongly suggests that ncRNAs to a larger extend may sculpt cell identity. Twelve ncRNAs correlated to Pou5f1, Sox2, and Nanog are identified, which may have functions during mouse embryonic stem cell (ES cell) differentiation. Interestingly, the promoter of two ncRNAs binding by Pou5f1 and Nanog are regulated by pluripotential factors to mediate chromatin reprogramming [51]. Even in retinal development, a 9 kb non-coding RNA 2 (RNCR2) targets to a specific nucleus domain [52], which helps regulate mouse retinal cell differentiation ncRNA are also important in regulating cell cycle. ncRNAs transcribed from multiple cyclin D1 promoter with the 200 and 330 bp length [30] were able to bind RNA-binding protein TLS and exert inhibition of p300 histone acetyltransferase activity to perform the transcriptional silencing. The further research [48] recognized that expression of promoter-associated ncRNAs is inducible in response to ionizing irradiation, which suggests that ncRNA can serve as a sensor of DNA damage. All these evidence show that ncRNA is a compelling element in various aspects from ES cell maintenance to differentiation, from cell cycle to DNA damage response.

ncRNA Generated and Targeted in Gene Body Regions Demanding More Consideration

Gene body region including the exon and intron have different patterns of chromatin modification from promoter region. Yet, similar to what happened in the promoter; enormous ncRNAs are generated and supposed to be biologically functional in gene body region. A combination of the coding transcripts along with the non-coding transcripts increases the transcriptional complexity. The *ab initio* reconstruction of transcriptome [53] reveals a large number of novel 5′start and/or 3′end transcripts and small RNAs. Relying on both computational and experimental assays, another study [54] identified that at least 4%–5% of the neighbour genes in the human genome could be transcribed into a single RNA unit. We hypothesize that those fusion long transcripts may function as ncRNA or be translated to chimeric protein contributing to the evolution of protein diversity. All these examples show that gene body regions are capable of producing unannoted ncRNAs whose functions need to be further determined.

The transcribed introns are actually the largest part of gene-body-derived ncRNA. A survey of mRNA and EST public database showed >55,000 ncRNAs derived from introns of 74% of all RefSeq genes. Interestingly, part of intronic ncRNAs and their corresponding mRNAs have affiliated expression pattern [55]. Whether intronic transcripts have function in regulating gene expression and modifying the chromatin status of gene body is still an open question. Despite all this, ∼25% of known human miRNAs are located in intronic regions [56]. The intronic miRNA, together with the host gene mRNA, synergistically regulate gene expression [57]. The intronic miRNAs are transcribed by Pol II or Pol III who prefer to initiate transcription through repetitive elements. Coincidently, many repeats often inhabit in the intronic region. In the human genome, repeat sequences that do not code for proteins make up at least 50% of the whole human genome, the precise function of such a large percentage of repeats remains to be explored. From an evolutionary historical view, repeats reshape the genome by rearranging it and thereby changing chromosome structure and dynamics [58]. Considering many Alu, making up ∼11% of the human genome [59] and primarily adjacent to GC-rich and gene-rich regions [60] are transcribed and processed further into siRNA, we proposed a model that endogenous Alu-derived siRNAs recruit RNA silencing complexes to the Alu-containing region to change chromatin accessibility thereby leading to the silence of genes surrounded by Alu during disease development.

Gene duplication generally occurs in evolutionary process contributing to genetic novelty. The extra copy would go through mutation, depletion, or fraction which brings about malfunctions to form a pseudogene. A series of genome-wide studies [61] identified ∼15,000 pseudogenes in human cell. Pseudogene is not, as its name suggests a counterfeit, but gains its function to regulate protein-coding transcripts similarity toncRNA. Some pseudogene transcripts serve as a shield to suppress mRNA degradation by RISC complex in which miRNA is complementary to both pseudogene transcript and mRNA [62] or other RNA-destabilizing factor [63], indicating the significance of pseudogene that can be seen as functional ncRNA.

Gene Terminal Region-derived ncRNAs and their Possible Roles in Gene Regulation

Gene terminal ncRNAs including antisense RNA may regulate gene expression in different ways. Interestingly, comparing with gene body, gene terminal region represented higher level of transcripts, which is similar to promoter region [22]. Furthermore, a potential processing into shorter-terminal-associated RNA has been investigated. TASRs and aTASR are two types of terminal-derived ncRNA. TASR was identified by tiling array [64] which share the same sequence with parental mRNA but shorter,

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024

and aTASR is a new class of small RNA with 5′-poly (U)-tailed emerging in gene terminal [18]. The biological functions of the termini-associated RNA need a thorough inquiry. Whether those terminal-derived ncRNAs are transcribed from the antisense strand or produced by taking 3′-poly(A)-tailed mRNA as template with the help of some factors like RNA-dependent RNA polymerase or RNAPII [65] needs more investigation. We hypothesize that it is similar to divergent transcription occurring in active promoter, TASRs may function like promoter-associated ncRNAs in regulating gene expression.

In addition to terminal-derived short ncRNA, the antisense RNA overlapping with the gene terminal transcribed from the downstream region in opposite directions can also be considered as a special type of terminal ncRNA. The prevalence of antisense transcripts is a common phenomenon in organisms from bacteria to mammalian [15,66,67]. The distribution of antisense RNAs implies their general participation in the regulation of gene expression with mechanisms related to chromatin modification. Through extensive base pairing, antisense RNA regulates sense transcripts more quickly and specifically by its nature. Although an antisense transcript that lies on 5′-UTR of Zeb2 indirectly activates Zeb2 expression via affecting Zeb2 mRNA splicing [68], it seems that actively transcriptional antisense RNA is exclusively involved in gene silencing. During *Saccharomyces cerevisiae* chronological aging [69], the antisense of phosphate transporter PHO84 assembled in Hda1/2/3 histone deacetylase complex leading to H3K18 deacetylation. P15 antisense can silence p15 transcription directly through increased H3K9 dimethylation and decreased H3K4 dimethylation, both of which represent gene silencing markers. The model reveals that antisense RNA mediates heterochromatin formation and then induces gene silencing. Afterwards, the changed chromatin architecture moreover leads to DNA methylation with a long-lasting heritable effect [39].

Intergenic ncRNA, from Junk to Wealth

The intergenic regions of genomic DNA have long been considered as junk DNA with less function, but an evolutionary comparison discovered astonishingly that intergenic transcripts constitute half of the expression differences between human and chimpanzee [70]. Recent studies have suggested that the mammalian genome transcribed several thousands of lincRNAs [9,28,53].

Many lincRNAs are associated with chromatin-modifying complexes such as PRC2, CoREST [31,53] and implicated in embryonic development, DNA damage repair, signal transmission, and cell cycle progression [9]. One lincRNA named lincRNA-p21, target of p53, holds fundamental duties in the p53 transcriptional response to apoptosis [71]. Another lincRNA, RNA-RoR, poses a positive impact to induce pluripotent stem cell [72], during which, RNA-RoR works as scaffold guiding chromatin-associated factors to particular loci to modulate chromatin state. Reprogramming, which refers to erasure and rewrite of epigenetic marks, emerges in many biological process. In human and mouse, there are large K9-modified regions that stretch up to 4.9 Mb with the help of H3K9methyltransferase G9a, to reorganize specific genome regions into heterochromatin [3]. We speculate that the lincRNA may recruit chromatin modification complex to contribute to the formation of large organized chromatin K9 modifications (LOCKs). More research [73] has found that close relationship between LOCK and chromatin remodeling during reprogramming of epithelial-to-mesenchymal transition. Searching for lincRNA that guide remodeling complexes to the large region in this transition process may benefit to illustrate the intrinsic mechanism. Some lincRNAs transcribed from a relatively large region may represent an obvious chromatin remodeling which in turn regulates its nearby gene activity. In contrast to repressive effect of lincRNA-p21 [71], two lincRNAs, ncRNA-a6, and ncRNA-a7 [74], located nearby Snail1 and Snail2, respectively, are positively correlated with Snail expression.

Hox genes are a cluster of related genes that establish the embryonic structure and orientation of an organism including HOXA, HOXB, HOXC, and HOXD. There are plentiful of polycomb/Trithorax group response elements (PRE/TREs) diffusely spreading across the genome [75], and ncRNAs with transcribed TRE sites can recruit trithorax complex, an indicator of active gene expression, to the corresponding genomic TRE loci to elevate gene expression in *cis* [76]. HOTTIP transcribed from the HOXA locus brings WDR5/MLL complex to active several HOXA genes expression [77].

Enhancer locates far upstream or downstream of the gene regulating gene expression either in *cis* or in *trans*. Mediator complex, which also interact with Pol II [78,79], usually enhances forming loop with the promoter to play a positive impact in gene expression from a long distance. Different from promoter, which prefers to control housekeeping gene, enhancer seems to regulate tissue-specific genes [7] consistent with more enhancer diversity than that of promoter in various cell types. Prominently, most disease-related single-nucleotide polymorphism (SNP) goes through a significant enrichment in the enhancer region; these SNPs probably affect the enhancer function by transcribing different levels of enhancer-related ncRNA. Kim *et al*. [80] described the identification of $>$12,000 enhancer-derived ncRNAs, namely eRNAs, in the mouse neuronal cell. The level of eRNA for each enhancer determined by RNA-Seq reads $\sim$1.5 kb region on both sides of the p300/CBP peak. In human prostatic cells, treated with

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024

androgen, enhancers responding to androgen receptor undergone a widespread transcriptional reprogramming [44]. These inducible eRNAs function as a linker between an activated enhancer and its affiliated promoter to regulate gene expression.

Other intergenic RNA including transcribed repeats especially presented in centromere and telomere to form heterochromatin play a profound role in maintaining genomic integrity. Several studies show that telomeric regions produce large telomere repeat-containing RNA and small RNAs [81,82] and centromeric repeats also undergo a prevalent bidirectional transcription [83]. Deletion of argonaute, a major component of RNAi machinery, results in the disruption of centromeric heterochromatin. Transient transfection GFP-siRNA leads to the formation of siRNA machinery surrounding telomeres, which indicates that telomeric repeat-containing RNA that serves as target or byproduct is always correlated with chromosome end [81].

Conclusion and Perspective

While the majority of the genome is actively transcribed into RNA, most of them belong to ncRNA scope. This decade, a series of studies show that long ncRNA (antisense, lincRNA) and short RNA (siRNA, piRNA, promoter-associated RNA) are correlated with chromatin-associated complex recruited to particular genome loci to change chromatin structure. It is still a long way for us to understand the intrinsic mechanism thoroughly in mammals. The discovery of the positive relationship between high-methylation level in gene body and actively transcriptional genes [84,85] and higher frequency of non-CpG methylation in exons than that in introns [86] need further investigation. Considering both exon and intron transcribed as the same unit, the difference in non-CpG methylation frequency may be attributed to the processed introns that serve as ncRNA in turn to recruit modification enzymes for chromatin remodeling. Above all, although obtaining great achievement in the field of RNA-directed chromatin modification, the underlying mechanism as a whole is elusive. ncRNA as a key player in remodeling epigenetic inheritance needs to be further investigated to understand gene expression and disease development.

Acknowledgements

The authors would like to thank Yue Yu, Xiao Yao, and Jin Li for critical reading of the manuscript. We also thank Yan Li for comments and Min Xiao for advice and drawing the figure.

Funding

This work was supported by the grant from the Ministry of Science and Technology of China (2009CB825603).

References

1 Zhang L, Ma H and Pugh BF. Stable and dynamic nucleosome states during a meiotic developmental process. Genome Res 2011, 21: 875–884.
2 Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W and Eussen BH, *et al*. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 2008, 453: 948–951.
3 Wen B, Wu H, Shinkai Y, Irizarry RA and Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat Genet 2009, 41: 246–250.
4 Fontan-Lozano A, Suarez-Pereira I, Horrillo A, Del-Pozo-Martin Y, Hmadcha A and Carrion AM. Histone H1 poly[ADP]-ribosylation regulates the chromatin alterations required for learning consolidation. J Neurosci 2010, 30: 13305–13313.
5 Choy JS, Wei S, Lee JY, Tan S, Chu S and Lee TH. DNA methylation increases nucleosome compaction and rigidity. J Am Chem Soc 2010, 132: 1782–1783.
6 Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M and McLoughlin EM, *et al*. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 2011, 473: 394–397.
7 Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB and Zhang X, *et al*. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 2011, 473: 43–49.
8 Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G and Alvarez P, *et al*. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 2007, 448: 553–560.
9 Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D and Huarte M, *et al*. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009, 458: 223–227.
10 Myer VE and Young RA. RNA polymerase II holoenzymes and subcomplexes. J Biol Chem 1998, 273: 27757–27760.
11 Hirota K, Miyoshi T, Kugou K, Hoffman CS, Shibata T and Ohta K. Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature 2008, 456: 130–134.
12 De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK and Muller H, *et al*. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol 2010, 8: e1000384.
13 Barski A, Chepelev I, Liko D, Cuddapah S, Fleming AB, Birch J and Cui K, *et al*. Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. Nat Struct Mol Biol 2010, 17: 629–634.
14 Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH and Kim VN. MicroRNA genes are transcribed by RNA polymerase II. EMBO J 2004, 23: 4051–4060.
15 Yang N and Kazazian HJ. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol 2006, 13: 763–771.
16 Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V and Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Science 2006, 313: 320–324.
17 Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA and Young RA, *et al*. Divergent transcription from active promoters. Science 2008, 322: 1849–1851.
18 Kapranov P, Ozsolak F, Kim SW, Foissac S, Lipson D, Hart C and Roels S, *et al*. New class of gene-termini-associated human RNAs suggests a novel RNA copying mechanism. Nature 2010, 466: 642–646.

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024

19 Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K and Levine M, *et al*. RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat Genet 2007, 39: 1512–1516.

20 Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ and Brow DA. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell 2006, 24: 735–746.

21 Nechaev S, Fargo DC, Dos SG, Liu L, Gao Y and Adelman K. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science 2010, 327: 335–338.

22 Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT and Kraus WL. A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell 2011, 145: 622–634.

23 Churchman LS and Weissman JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 2011, 469: 368–373.

24 Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH and Weng Z, *et al*. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007, 447: 799–816.

25 Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S and Long J, *et al*. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science 2005, 308: 1149–1154.

26 Han J, Kim D and Morris KV. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. Proc Natl Acad Sci USA 2007, 104: 12422–12427.

27 Martianov I, Ramadass A, Serra BA, Chow N and Akoulitchev A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature 2007, 445: 666–670.

28 Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea MD, Thomas K and Presser 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.

29 Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N and Oyama, *et al*. The transcriptional landscape of the mammalian genome. Science 2005, 309: 1559–1563.

30 Wang X, Arai S, Song X, Reichart D, Du K and Pascual G, *et al*. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature 2008, 454: 126–130.

31 Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F and Shi Y, *et al*. Long noncoding RNA as modular scaffold of histone modification complexes. Science 2010, 329: 689–693.

32 Vu TH, Chuyen NV, Li T and Hoffman AR. Loss of imprinting of IGF2 sense and antisense transcripts in Wilms' tumor. Cancer Res 2003, 63: 1900–1905.

33 Vu TH and Hoffman AR. Promoter-specific imprinting of the human insulin-like growth factor-II gene. Nature 1994, 371: 714–717.

34 Sleutels F, Zwart R and Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature 2002, 415: 810–813.

35 Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC and Feil R, *et al*. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science 2008, 322: 1717–1720.

36 Lee JT, Lu N and Han Y. Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain. Proc Natl Acad Sci USA 1999, 96: 3836–3841.

37 Zhao J, Sun BK, Erwin JA, Song JJ and Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008, 322: 750–756.

38 Kim DH, Villeneuve LM, Morris KV and Rossi JJ. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat Struct Mol Biol 2006, 13: 793–797.

39 Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP and Cui H. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature 2008, 451: 202–206.

40 Kawasaki H and Taira K. Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature 2004, 431: 211–217.

41 Morris KV, Chan SW, Jacobsen SE and Looney DJ. Small interfering RNA-induced transcriptional gene silencing in human cells. Science 2004, 305: 1289–1292.

42 Watanabe T, Tomizawa S, Mitsuya K, Totoki Y, Yamamoto Y, Kuramochi-Miyagawa S and Iida N, *et al*. Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. Science 2011, 332: 848–852.

43 Chan SW. Inputs and outputs for chromatin-targeted RNAi. Trends Plant Sci 2008, 13: 383–389.

44 Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y and Qiu J, *et al*. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 2011, 474: 390–394.

45 Morris KV, Santoso S, Turner AM, Pastori C and Hawkins PG. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. PLoS Genet 2008, 4: e1000258.

46 Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI and Moazed D. RNAi-mediated targeting of heterochromatin by the RITS complex. Science 2004, 303: 672–676.

47 Yue X, Schwartz JC, Chu Y, Younger ST, Gagnon KT and Elbashir S, *et al*. Transcriptional regulation by small RNAs at sequences downstream from 3′ gene termini. Nat Chem Biol 2010, 6: 621–629.

48 Place RF, Li LC, Pookot D, Noonan EJ and Dahiya R. MicroRNA-373 induces expression of genes with complementary promoter sequences. Proc Natl Acad Sci USA 2008, 105: 1608–1613.

49 Napoli S, Pastori C, Magistri M, Carbone GM and Catapano CV. Promoter-specific transcriptional interference and c-myc gene silencing by siRNAs in human cells. EMBO J 2009, 28: 1708–1719.

50 Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD and Horlings HM, *et al*. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet 2011, 43: 621–629.

51 Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB and Askarian-Amiri ME, *et al*. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res 2008, 18: 1433–1445.

52 Rapicavoli NA, Poth EM and Blackshaw S. The long noncoding RNA RNCR2 directs mouse retinal cell specification. BMC Dev Biol 2010, 10: 49.

53 Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X and Fan L, *et al*. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol 2010, 28: 503–510.

54 Parra G, Reymond A, Dabbouseh N, Dermitzakis ET, Castelo R, Thomson TM and Antonarakis SE, *et al*. Tandem chimerism as a means to increase protein complexity in the human genome. Genome Res 2006, 16: 37–44.

55 Nakaya HI, Amaral PP, Louro R, Lopes A, Fachel AA, Moreira YB and El-Jundi TA, *et al*. Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. Genome Biol 2007, 8: R43.

56 Li SC, Tang P and Lin WC. Intronic microRNA: discovery and biological implications. DNA Cell Biol 2007, 26: 195–207.

57 Barik S. An intronic microRNA silences genes that are functionally antagonistic to its host gene. Nucleic Acids Res 2008, 36: 5232–5241.

58 Hancock JM and Simon M. Simple sequence repeats in proteins and their significance for network evolution. Gene 2005, 345: 113–118.

59 Cordaux R and Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet 2009, 10: 691–703.

60 Korenberg JR and Rykowski MC. Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. Cell 1988, 53: 391–400.

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024

61 Bischof JM, Chiang AP, Scheetz TE, Stone EM, Casavant TL, Sheffield VC and Braun TA. Genome-wide identification of pseudogenes capable of disease-causing gene conversion. Hum Mutat 2006, 27: 545–552.
62 Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ and Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature 2010, 465: 1033–1038.
63 Hirotsune S, Yoshida N, Chen A, Garrett L, Sugiyama F, Takahashi S, Yagami K and Yoshiki A., *et al*. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 2003, 423: 91–96.
64 Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT and Stadler PF, *et al*. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 2007, 316: 1484–1488.
65 Lehmann E, Brueckner F and Cramer P. Molecular basis of RNA-dependent RNA polymerase II activity. Nature 2007, 450: 445–449.
66 Dornenburg JE, Devita AM, Palumbo MJ and Wade JT. Widespread antisense transcription in Escherichia coli. MBio 2010, 1: e00024–10.
67 Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M and Nishida H, *et al*. Antisense transcription in the mammalian transcriptome. Science 2005, 309: 1564–1566.
68 Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB and Pena R, *et al*. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev 2008, 22: 756–769.
69 Camblong J, Iglesias N, Fickentscher C, Dieppois G and Stutz F. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell 2007, 131: 706–717.
70 Khaitovich P, Kelso J, Franz H, Visagie J, Giger T, Joerchel S and Petzold E, *et al*. Functionality of intergenic transcription: an evolutionary comparison. PLoS Genet 2006, 2: e171.
71 Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D and Khalil AM, *et al*. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell 2010, 142: 409–419.
72 Loewer S, Cabili MN, Guttman M, Loh YH, Thomas K, Park IH and Garber M, *et al*. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat Genet 2010, 42: 1113–1117.
73 McDonald OG, Wu H, Timp W, Doi A and Feinberg AP. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. Nat Struct Mol Biol 2011, 18: 867–874.
74 Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G and Lai F, *et al*. Long noncoding RNAs with enhancer-like function in human cells. Cell 2010, 143: 46–58.
75 Ringrose L and Paro R. Polycomb/trithorax response elements and epigenetic memory of cell identity. Development 2007, 134: 223–232.
76 Sanchez-Elsner T, Gou D, Kremmer E and Sauer F. Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to ultrabithorax. Science 2006, 311: 1118–1123.
77 Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R and Chen Y, *et al*. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 2011, 472: 120–124.
78 Taatjes DJ. The human Mediator complex: a versatile, genome-wide regulator of transcription. Trends Biochem Sci 2010, 35: 315–322.
79 Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL and Ebmeier CC, *et al*. Mediator and cohesin connect gene expression and chromatin architecture. Nature 2010, 467: 430–435.
80 Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J and Harmin DA, *et al*. Widespread transcription at neuronal activity-regulated enhancers. Nature 2010, 465: 182–187.
81 Ho CY, Murnane JP, Yeung AK, Ng HK and Lo AW. Telomeres acquire distinct heterochromatin characteristics during siRNA-induced RNA interference in mouse cells. Curr Biol 2008, 18: 183–187.
82 Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E and Lingner J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science 2007, 318: 798–801.
83 Halic M and Moazed D. Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. Cell 2010, 140: 504–516.
84 Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH and Xie B, *et al*. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 2009, 27: 361–368.
85 Hellman A and Chess A. Gene body-specific methylation on the active X chromosome. Science 2007, 315: 1141–1143.
86 Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J and Nery JR, *et al*. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009, 462: 315–322.

Downloaded from https://academic.oup.com/abbs/article/44/1/40/1164 by guest on 24 April 2024