

Published in final edited form as:

*Nat Rev Mol Cell Biol.* 2009 September ; 10(9): 637–643. doi:10.1038/nrm2738.

## Regulatory roles of natural antisense transcripts

Mohammad Ali Faghihi and Claes Wahlestedt\*

Molecular and Integrative Neurosciences Department (MIND), The Scripps Research Institute, Jupiter, FL, 33458, USA

### Abstract

Mammalian genomes encode numerous natural antisense transcripts, but the function of these transcripts is not well understood. Functional validation studies indicate that antisense transcripts are not a uniform regulatory RNA group, but instead belong to multiple categories with some common features. Recent evidence indicates that antisense transcripts are frequently functional and use diverse transcriptional and post-transcriptional gene regulatory mechanisms to carry out a wide variety of biological roles.

---

Natural antisense transcripts are RNA molecules that are transcribed from the opposite DNA strand compared with other transcripts and overlap in part with sense RNA. Both sense and antisense RNAs can encode proteins or be non-protein-coding transcripts; however, the most prominent form of antisense transcription in the mammalian genome is a non-protein-coding antisense RNA partner of a protein-coding gene<sup>1</sup>. The presence of non-protein-coding sense and antisense transcript pairs implies that natural antisense transcripts may also regulate non-protein-coding sense RNAs.

Various techniques have been used to identify natural antisense transcripts, including large-scale sequencing of cDNA clones<sup>1</sup>, tiling arrays<sup>2</sup>, bioinformatics analysis of RefSeq and EST databases<sup>3</sup>, hybridization techniques<sup>4</sup>, SAGE libraries<sup>5</sup>, strand-specific microarrays<sup>6</sup> and, most recently, asymmetric strand-specific analysis of gene expression (ASSAGE)<sup>7</sup> and global run-on sequencing (GRO-seq)<sup>8</sup>. These studies have demonstrated widespread antisense transcription in mammalian genomes. The mammalian transcriptomes are still not sequenced deeply enough to provide the organization of all low copy number antisense transcripts. Nevertheless, the majority of transcription units (TUs; see Glossary box) appear to contain antisense transcripts<sup>1</sup>. Moreover, antisense transcripts have been **documented for** many active promoters<sup>8-10</sup>. Therefore, the presence of antisense transcripts is no longer a curiosity but rather a pervasive feature of mammalian genomes. For instance, the FANTOM-3 mouse transcriptome sequencing consortium identified natural antisense transcripts for more than 70% of TUs, the majority of which represent non-protein-coding RNA<sup>1</sup>.

### Characteristics of antisense RNAs

Antisense transcripts are not evenly distributed across the genome. Both ends of protein-coding genes have a propensity for natural antisense transcription<sup>7,11</sup>; specifically, antisense transcription is enriched 250 nucleotides upstream of the transcription start site (TSS)<sup>8,9</sup> and 1.5 kilobases downstream of sense genes<sup>8,12</sup>. The basal expression levels of sense and antisense transcripts in different tissues and cell lines may be either positively or negatively correlated<sup>1,13</sup>. Moreover, antisense transcript expression in different human cell lines is not

---

\*Correspondence should be addressed to: Dr. Claes Wahlestedt, The Scripps Research Institute, 130 Scripps Way, Jupiter, FL, 33458, USA, clawah@scripps.edu, Telephone: +1 561 228 3462, Fax: +1 561 228 3079.

always linked to expression of the sense gene, which suggests the use of alternative transcriptional regulatory elements<sup>7</sup>. Antisense RNAs have a tendency to undergo fewer splicing events and typically show lower abundance compared with sense transcripts<sup>7</sup>. Interestingly, knockdown or blockade of endogenous antisense transcripts can have multiple outcomes, with the corresponding sense transcript concentration showing either an increase (discordant regulation) or a decrease (concordant regulation)<sup>14</sup>. It has been proposed that discordant de-repression of sense transcript expression, resulting in upregulation of sense RNA expression, can be achieved by removal or steric blockade of many but not all antisense transcripts<sup>1,14</sup>. These and other variable intrinsic properties imply that antisense-mediated regulation of gene expression must operate through a variety of mechanisms, and further suggest that antisense transcripts are a heterogeneous group of regulatory RNAs.

Supplementary information S1 (table) contains all reported functional antisense transcripts in mammalian genomes characterized to date. This growing list of validated sense–antisense transcript pairs includes many important developmental genes, as well as genes known to be involved in a variety of human disorders. Natural antisense transcripts must therefore be assumed to exert positive or negative regulation at different levels of gene expression invoking diverse and largely uncharacterized co-regulatory factors. Hence, potential involvement of natural antisense transcripts should be considered in a wide range of future studies on the regulation of gene expression.

There are a number of proposed mechanisms for antisense-mediated regulation of sense mRNA. Here, we categorize the mechanisms into four main groups: mechanisms related to transcription, RNA–DNA interactions, RNA–RNA interactions in the nucleus and RNA–RNA interactions in the cytoplasm. Each type of mechanism will be discussed in detail below.

## Transcription-related modulation

According to this mechanism, the act of transcription in the antisense direction, but not the antisense RNA molecule itself, modulates transcription of sense RNA. Transcription in the antisense direction is suggested to cause alterations in sense RNA, as in the case of transcriptional collision, and is also proposed to be involved in genomic rearrangements (Figure 1).

### Transcriptional collision

The transcriptional collision model is based on the assumption that RNA polymerases bind to the promoters of convergent genes on opposing strands of DNA and then proceed towards the 3' end of sense and antisense genes. RNA polymerase complexes collide in the overlapping region, blocking further transcription<sup>15</sup> (Figure 1a). Collision of RNA polymerases has been observed in *Escherichia coli*, using atomic force microscopy<sup>16</sup>. Transcriptional collision has also been observed in *Saccharomyces cerevisiae*<sup>17</sup>, where transcription from the antisense direction halts transcription of the sense mRNA. A bioinformatics study found that, in humans and mice, the length of overlapping regions between sense and antisense transcripts is inversely correlated to the expression levels of antisense transcript. This finding may suggest a possible clash of RNA polymerases, pertaining to the point that the likelihood of collision increases as the length of overlapping region increases<sup>18</sup>. In this model, the antisense RNA *per se* is not involved in interfering with transcription of the sense RNA: instead, it is transcription in the antisense direction that decreases synthesis of sense RNA.

Transcriptional collision may affect a subset of mammalian *cis*-encoded natural antisense transcripts, but it is apparently not the predominant route of antisense-mediated gene regulation. Transcription of the sense and antisense RNAs may occur at different times at the same chromosome locus, or sense and antisense RNA may be transcribed simultaneously from

paternal and maternal chromosomes. Allelic-specific transcription may explain why the X chromosome shows a significantly lower degree of antisense transcription compared with other chromosomes<sup>1</sup>.

### Genomic rearrangements

Immunoglobulin production in B lymphocytes and receptor selection by T lymphocytes depend on transcription from hypervariable regions. To generate variability, T and B lymphocytes need a recombination process that occurs through hypermutation in the variable regions of immunoglobulin and T cell receptor genes. Activation-induced cytidine deaminase (AID), which deaminates deoxycytidine to deoxyuridine in single-stranded DNA (ssDNA), is required for both somatic hypermutation (SHM) and class switch recombination (CSR), two processes that are involved in secondary immune response. DNA must locally uncoil or melt, and become single-stranded for DNA recombination, replication and transcription to occur. The transcription complex consists of an RNA polymerase molecule enclosing approximately 17 +/- 1 melted bases (the transcription bubble) of the template DNA<sup>19</sup>; however, the size of transcription bubble may vary in different instances. The AID enzyme optimally deaminates a small ssDNA bubble, *in vitro*, consisting of ~5 nucleotides with a ssDNA motif, called WRC<sup>20</sup>. It has been hypothesized that antisense transcription in the variable region could make the ssDNA accessible for AID<sup>21,22</sup> (Figure 1b).

Antisense transcription is frequently observed in the variable<sup>23,24</sup>, but not in the constant, regions and it may be initiated in preparation for SHM and CSR (reviewed in Ref. 25). According to this model, antisense transcription is proposed to open the transcription bubble and expose the ssDNA, which is a substrate for AID<sup>23,24</sup>, or antisense transcription may be involved in remodelling of the chromatin structure to make the DNA sequence accessible for recombination<sup>22-24</sup>. In both of these scenarios, transcription from the antisense direction, but not the antisense RNA molecule, is important for genomic rearrangement.

### RNA–DNA interactions

Natural antisense transcripts may also be involved in epigenetic regulation of transcription, through DNA methylation<sup>26</sup>, chromatin modifications<sup>27</sup> and monoallelic expression (for example, genomic imprinting<sup>28</sup>, X-chromosome inactivation<sup>29</sup> and random monoallelic exclusion of the autosomal loci). This gene regulatory model is based on direct or indirect RNA–DNA and RNA–chromatin interactions<sup>30</sup>. Antisense transcripts may bind to the corresponding DNA strand, resulting in either DNA methylation (Figure-2a) or providing a scaffold for histone-modifying enzyme (HME) recruitment and subsequent alterations of the chromatin status (Figure-2b).

Some of the antisense-mediated epigenetic changes are reported to be independent of Dicer, arguing against a role for small RNA mediators<sup>27</sup>. Therefore, antisense RNA processing to small RNA, which may be mediated by Dicer in the cytoplasm, is not necessarily required for epigenetic modifications. Instead, antisense transcripts accumulate locally and trigger DNA or chromatin modifications<sup>31</sup>. These modifications then expand to neighbouring regions, even though these adjacent regions do not exhibit complementarities to the original antisense transcript<sup>32</sup>. The secondary expansion of the modifications may be restricted to the promoter or enhancer of a single gene (random monoallelic exclusion), or may include a cluster of genes, such as in the case of genomic imprinting of the *Kcnq1* imprinted locus<sup>28</sup>. Finally, this expansion process occasionally involves an entire chromosome, such as X chromosome inactivation in females.

## Alteration of DNA and chromatin

Natural antisense transcripts have been proposed to cause DNA methylation<sup>26</sup>, DNA demethylation<sup>33</sup> and chromatin modifications of non-imprinted autosomal loci<sup>27,34</sup>. Suppression of transcription is usually caused by DNA and chromatin modification at the promoter region of the sense strand (Supplementary information S2 (box)). For example, an antisense RNA for the  $\alpha$ -globulin 2 gene (*HBA2*) can induce DNA methylation, leading to silencing of the  $\alpha$ -globulin 2 gene<sup>26</sup>.

Antisense-mediated transcriptional silencing also affects the *p15*<sup>27</sup>, *p21*<sup>34</sup> and progesterone receptor (*PR*)<sup>35</sup> genes, through DNA methylation and heterochromatin formation. Suppression of sense transcription is often induced by trimethylation of histone H3 lysine 27 (H3K27me3) at the sense promoter region. For example, an antisense transcript of the tumour suppressor gene *p21* recruits a regulatory complex that induces H3K27me3 and suppression of the sense promoter region<sup>34</sup>. Considering the presence of RNA-binding motifs in some chromatin-modifying enzymes<sup>30</sup>, one could postulate that RNA transcripts are local modulators of chromatin structure. This proposed mechanism might also explain functionality even if the abundance of antisense RNA molecule is low<sup>9,10,34</sup>. Unlike RNAs that can be present in cells in many copies, there are two copies of DNA for any given gene in cells; therefore, only two molecules of antisense RNA per cell are sufficient to bind to the corresponding DNA strand and to exert a regulatory function. These proposed local effects of antisense transcripts might sometimes extend to the neighbouring regions, such as in the case of the imprinted locus<sup>28</sup> (see below). These examples suggest the possibility of local and extended antisense RNA-induced DNA and chromatin modifications.

## Chromatin alteration of active promoters

Overlapping transcription of small (<50 nucleotides) RNAs<sup>36,37</sup> as well as promoter-associated small RNAs (PASRs)<sup>36,37</sup>, termini-associated small RNAs (TASRs)<sup>36,37</sup> and promoter upstream transcripts (PROMPTs)<sup>10</sup> have been documented using strand-specific genomic tiling arrays in the ENCODE region of the human genome<sup>36,37</sup>. Transcription start site-associated RNAs (TSSa-RNAs) (20-90 nucleotides) have also been identified using deep sequencing techniques<sup>8,9</sup>. The enrichment of antisense transcripts in both promoter and terminal regions has been confirmed by the unbiased ASSAGE technique<sup>7</sup>. TSSa-RNAs are flanking active promoters in both the sense and antisense direction with regard to downstream protein-coding genes<sup>8,9</sup>. Divergent promoter activity, which is believed to produce this small RNA group, is documented for more than half of mouse and human genes<sup>8,9</sup>. These well-documented promoter activities, which generate promoter-directed sense and antisense transcripts, challenge the canonical 'gene' definition and how genes and their regulatory elements are arranged in mammalian genomes. Mammalian transcriptomes appear to contain an abundance of nested and overlapping gene structures, giving rise to both coding and noncoding transcripts. In addition, it seems that the RNA regulatory elements that control the expression of a gene are frequently distributed within or beyond other genes in both directions. Collectively, pervasive divergent promoter activity suggests a lack of definite 5' and 3' boundaries for the transcribed genes.

Interestingly, these small antisense transcripts do not correspond to annotated natural antisense transcripts<sup>36,37</sup>. There is also no evidence for double-stranded RNA (dsRNA) or hairpin RNA precursors that might represent intermediates in the biogenesis of such small natural antisense transcripts<sup>36,37</sup>. Generation of TSSa-RNA is also shown to be independent of Dicer, suggesting that the possible regulatory function of these transcripts is not mediated through common RNA interference (RNAi), such as RISC and RITS pathways. Additionally, the expression of promoter-directed antisense transcripts may be associated with the presence of abortive RNA transcripts<sup>38</sup>.

The transcribed regions of PROMPTs and TSSa-RNAs overlap with paused RNA polymerase II (RNAPII) and active chromatin marks, such as trimethylation of histone H3 lysine 4 (H3K4me3), suggesting a role for local RNA accumulation in maintaining the dynamic chromatin state required for promoter activity<sup>9,10</sup>. The abundance of short RNA transcripts at the transcription initiation and termination sites<sup>8,9</sup> as well as transcription of unstable transcripts upstream of genes<sup>10</sup>, which are usually co-localized with particular chromatin marks, imply involvement of these local and transient RNA transcripts in the regulation of the sense gene expression and possible participation in chromatin remodelling. We hypothesize that the frequent presence of these RNA classes in or around the promoter region of actively transcribed genes, in a very low concentration, suggests the possibility of local RNA-induced chromatin modifications.

### Genomic imprinting

Imprinted genes are genes for which only one allele, maternal or paternal, is actively transcribed. Natural antisense transcripts are often associated with imprinted genes, with a frequency up to 81% in one study<sup>1</sup> (Supplementary information S3 (box)). More than 160 imprinted genes have been identified so far in humans and mice, most of which are organized into clusters. Some imprinted genes, including insulin-like growth-factor type-2 receptor (*IGFER*)<sup>39</sup> and potassium voltage-gated channel, KQT-like (*Kcnq1*) imprinting control region<sup>31</sup>, exhibit guided chromatin and DNA modification by antisense RNA that expands to include neighbouring genes. These effects are not mediated through the RNAi pathway<sup>40</sup>. Instead, antisense RNA appears to recruit repressor complexes that modify chromatin into an inactive state. Suppressive chromatin modifications spread in both directions to neighbouring genes, similar to X chromosome inactivation (below) but with limited penetrance<sup>31</sup>.

### X chromosome inactivation

X chromosome inactivation is required for balanced expression of the genes on the X chromosome in female mammals. This dosage compensation occurs through heterochromatin formation along the X chromosome to be inactivated. Two long non-protein-coding genes are transcribed from the X chromosome inactivation centre, *XIST* (X-inactive specific transcript) and *TSIX* (X [inactive]-specific transcript, antisense). These two sense and antisense transcripts control the silencing of the X chromosome<sup>29</sup>. The X chromosome inactivation centre is necessary and sufficient for X chromosome inactivation. Elements of the mechanism have been demonstrated by showing that *Tsix* silences *Xist* through modification of the chromatin structure in the *Xist* promoter region. Premature termination of *Tsix* transcription abolishes the repressive chromatin configuration at the *Xist* promoter<sup>29</sup>.

### Nuclear RNA duplex formation

The third type of mechanism for antisense transcript-mediated gene regulation is based on the predicted nuclear sense–antisense RNA duplex formation (Figure 3Aa). The duplex RNA between sense and antisense transcripts may result in several outcomes, all of which modulate sense mRNA expression. Nuclear RNA duplex may produce mRNA transcripts with alternative splicing. Additionally, hybrid nuclear RNA is a substrate for editing enzymes that can alter the localization, transport and stability of the sense mRNA transcript.

### Alternative splicing and termination

Antisense RNA may also bind to the sense RNA, masking the splice sites and thereby changing the balances between splice variants (Figure 3Ab). Thyroid hormone receptor alpha gene (*TRa*) is an example where the antisense transcript *RevErbAa* influences splicing of *TRa1* and *TRa2* mRNAs<sup>41</sup>. Using a similar arrangement, antisense transcripts can potentially cause alternative polyadenylation and termination of transcription.

## Transport, nuclear retention and editing

Natural antisense transcripts may modulate mRNA nuclear transport by a mechanism that involves duplex formation between sense and antisense RNAs. Nuclear retention of the antisense RNA is commonly observed and may account for some antisense RNA-mediated regulation. Some cellular stressors, such as hypoxia, serum starvation and hydrogen peroxide, can change the nuclear retention pattern of antisense transcripts, thereby altering the levels of their sense partners<sup>42</sup>. Direct interactions with nuclear proteins or other nuclear RNAs likely cause nuclear retention of antisense transcripts, although the precise mechanism for each case has not yet been discovered.

Natural antisense transcripts have also been linked to mRNA editing (Figure 3Ac). Interaction between the *Drosophila melanogaster 4f-rnp* gene and its antisense transcript, *sas-10*, is reported to induce adenosine-to-inosine (A-to-I) editing in the overlapping region of *4f-rnp* mRNA<sup>43</sup>. A-to-I RNA editing is induced by dsRNA formation, resulting in the recruitment of the enzyme ADAR (adenosine deaminases that act on RNA), which deaminates the targeted adenosine to inosine<sup>44</sup>.

## Cytoplasmic RNA duplex formation

In the fourth proposed mechanism of sense–antisense interference, a duplex forms between sense and antisense RNA in the cytoplasm (Figure 3Ba–c). Cytoplasmic RNA hairpins may affect sense mRNA stability or translation. The sense–antisense RNA duplex may also, in theory, cover microRNA (miRNA)-binding sites or serve as a hairpin template for generating endogenous small interfering RNA (siRNA).

## Changes in mRNA stability and translation

Cytoplasmic sense–antisense duplex formation can alter sense mRNA stability and translation efficiency. The overlapping region may affect mRNA stability by reducing mRNA decay, whereby mRNA undergoes endo- or exonucleolytic degradation by various RNases. Indeed, we have recently demonstrated that an antisense *BACE1* transcript (*BACE1-AS*) increases the stability of *BACE1* mRNA through a mechanism involving RNA duplex formation. We hypothesize that transient RNA duplex formation may alter the secondary or tertiary structure of *BACE1* and thereby increase its stability<sup>42</sup>.

Antisense transcripts for inducible nitric oxide synthase (*iNOS*), an important gene in inflammatory diseases, increase the stability of *iNOS* mRNA<sup>45</sup>. Enhancement of *iNOS* mRNA stability is mediated through interactions of antisense RNA molecule with the AU-rich element (ARE)-binding HuR protein. The HuR protein, in turn, may suppress RNA degradation by inhibiting deadenylase or exonuclease enzymes<sup>45</sup>. Alterations in the secondary structure of HIF-1 $\alpha$  sense transcript by an antisense RNA may expose the ARE and reduce the stability of the transcript by making this RNA prone to degradation<sup>46</sup>.

Translational inhibition is yet another proposed function for some naturally occurring antisense, as in the case of B cell maturation antigen (*BCMA*) transcript, where over-expression of the antisense RNA has been reported to reduce sense protein, but not sense mRNA, levels<sup>47</sup>. Another well-documented case of translational inhibition is the antisense for *PU.1* mRNA. *PU.1* mRNA translation is inhibited by a noncoding antisense transcript<sup>48</sup>, which is a polyadenylated RNA with a lower concentration but a longer half-life time than the sense *PU.1* transcript. Processed antisense RNA in the cytoplasm may bind to the sense transcript and stall translation between initiation and elongation steps<sup>48</sup>. It also seems likely that formation of duplexes may alter ribosome entry, and consequently the protein output of the target.

## Masking miRNA-binding sites

We propose that many natural antisense transcripts may have the ability to cover miRNA-binding sites following cytosolic RNA duplex formation. This appears to be the case for an antisense *BACE1* transcript (*BACE1-AS*) in addition to its ability to increase the stability of *BACE1* mRNA<sup>42</sup>. Thus, one of the regulatory functions of antisense transcripts may be to 'mask' the miRNA-binding site on the sense mRNA. The 3'UTR mRNAs are frequently shown to contain target sites for miRNAs, and at least 34% of natural antisense transcripts in the FANTOM3 dataset showed tail-to-tail format with a 3'UTR overlapping region.

## Formation of endogenous siRNA

The presence of endogenous processing machinery for exogenous siRNA, which mediate sequence-specific knockdown of targeted genes, implies that endogenous siRNA should exist. Endogenous siRNAs derived from a natural antisense transcripts were observed in *Arabidopsis thaliana*, where they regulate salt tolerance<sup>49</sup>. Plant endogenous siRNA are derived from sense-antisense RNA duplex formation of several genes: for example, the *Sho* gene antisense transcript<sup>50</sup>, *SRLK/AtRAP* antisense transcript<sup>51</sup> and 64% of protein-coding natural antisense transcripts in *A. thaliana* are reported to generate endogenous siRNAs<sup>52</sup>.

Endogenous siRNAs derived from transposable elements and pseudogenes have also been identified in mouse oocytes and cultured human cells<sup>53-57</sup>. Endogenous siRNAs originating from mRNAs and their corresponding antisense were also recently identified in mouse oocytes<sup>54</sup> and human HepG2 liver carcinoma cells<sup>55</sup>. Both Piwi-interacting RNAs (piRNAs)<sup>58,59</sup> and siRNAs originating from mRNAs were found in mouse oocytes<sup>54</sup>. Therefore, formation of endogenous siRNA from naturally occurring antisense transcripts appears to occur in mammalian cells.

Transposable elements, inverted repeat structures, sense-antisense genes (*cis*-natural antisense transcript) and antisense transcripts from remote loci (*trans*-natural antisense transcripts) have also been recognized as sources of dsRNAs and subsequent, Dicer-dependent, endogenous siRNA production<sup>54,55</sup>. One example includes the overlapping transcripts in the kinesin family member 4A (*KIF4A*) and PDZ domain-containing 11 (*Pdzd11*) locus that generates endogenous siRNA derived from its *cis*-antisense transcript. Importantly, almost all of the endogenous siRNAs (117 unique sequences) are derived from the overlapping region of the sense and antisense RNA, suggesting that these endogenous siRNAs are produced from an intermolecular dsRNA formed between the oppositely oriented transcripts. In Dicer mutants, levels of the siRNAs derived from the *Pdzd11/Kif4* locus are decreased, and *Pdzd11* and *Kif4* mRNA levels are increased, suggesting that endogenous siRNA production is cytoplasmic and Dicer dependent, and that expression of both sense and antisense RNAs is regulated by the RNAi pathway<sup>54</sup>. Therefore, endogenous siRNAs may regulate both sense and antisense transcript levels.

Co-expression of natural antisense transcripts with their sense counterparts, as well as frequently observed concordant regulation of sense and antisense RNAs in many tissues and cell lines, argues against endogenous siRNA being the sole mechanism of antisense-mediated regulation of gene expression. Additionally, many co-expressed sense-antisense transcripts in *D. melanogaster* S2 cells do not generate endogenous siRNAs<sup>60,61</sup>. It is unclear how the majority of co-expressed sense and antisense RNAs escape the endogenous siRNA formation pathway<sup>63</sup>. It is also not clear whether there is active selection for entry into the RNAi pathway and endogenous siRNA formation.

## Concluding remarks

We have presented examples of functional natural antisense transcripts in order to show the multilayered involvement of these molecules in regulating gene expression. It can be concluded that antisense transcripts are not a uniform group of regulatory elements, but that they display some common features, including sequence complementarities to conventional sense genes. We have summarized the proposed mechanisms of antisense transcript actions into four models, each corresponding to a putative regulatory mechanism for certain antisense transcripts.

Although we have provided examples for each model, it is important to note that not all four models have equally strong experimental support. For instance, the frequently observed convergent promoter activity and engagement of multiple RNAPII complexes in the opposite orientation of active promoters would argue against the transcriptional collision model as a widely used mode of action. On the other hand, nuclear and cytoplasmic RNA–RNA interactions with various outcomes, such as RNA editing, splicing and dicing, have been observed for several functionally validated antisense transcripts, but it is not a predominant basis of natural antisense regulation. Furthermore, antisense RNA-induced alteration in sense promoter DNA methylation has only a few documented examples and, considering the static nature of DNA methylation, we postulate that it might be operational only in certain early developmental stages. Conversely, antisense RNA-induced chromatin remodelling seems to be a more feasible and dynamic mode of action for many low-copy natural antisense transcripts. In the latter case, antisense RNA might predominantly act locally to maintain or modify chromatin structure and ultimately to activate or suppress sense gene expression. Natural antisense transcripts are involved in different gene regulatory pathways, but it is still not clear which intrinsic properties of natural antisense RNA molecules, or extrinsic features such as protein interactions, cellular and developmental context, are decisive for any given pathway.

It is becoming clear that the expression of natural antisense transcripts and other non-protein-coding RNAs are pervasive in the human genome, although the mechanisms of their production and preferential sites of action are poorly understood. Particularly, the exact molecular machineries behind the antisense RNA-induced chromatin modifications are not currently known. It will be important to learn which histone-modifying enzymes might be involved, how they interact with antisense RNAs and which histone modifications they induce over time. It will also be of interest to define the exact contexts in which perturbation of the same antisense transcript<sup>35</sup> results in either activation or repression of sense RNA expression. It is remarkable that these fundamental and powerful regulatory mechanisms relating to natural antisense transcription remain to be elucidated in large part. It can be safely assumed that examination of natural antisense phenomena will remain a rich field of investigation for the foreseeable future.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We are grateful to M. P. van der Brug, B. H. Miller, J. Silva and S. Brothers for insightful comments and careful reading of the manuscript. Discussions with Y. Hayashizaki, G. St-Laurent and other colleagues within the FANTOM Transcriptomics consortium have also been highly valuable to us.



## Glossary box

Transcription unit	A group of expressed sequence tags or mRNAs, usually with alternative splice patterns, in which they share exonic overlap of at least one nucleotide and are in the same chromosomal orientation
Dicer	An RNase III family endonuclease that processes double-stranded RNA and pre-microRNAs into small interfering RNAs and microRNAs, respectively
ENCODE	The ENCyclopedia Of DNA Elements is a publicly founded project that aims to find functional elements in human genome ( <a href="http://www.genome.gov/10005107#1">http://www.genome.gov/10005107#1</a> )
MicroRNA	A small (20–25 nucleotides long) single-stranded RNA that is thought to regulate the expression of other genes, either through inhibiting protein translation or through degrading a target mRNA transcript, by a process that is similar to RNA interference
Short interfering RNA	A short (21–23 nucleotides long) RNA molecule that is processed from a long double-stranded RNA. siRNAs are functional components of the RNAi-induced silencing complex, (RISC) and typically target mRNAs by binding perfectly complementary sequences in the mRNA and causing their degradation
RISC	RNA-induced silencing complex is a multi-protein complex that incorporates one strand of siRNA and use them to recognize the target mRNA for degradation
RITS	RNA-induced transcriptional silencing complex is a multi-protein complex, e.g. in fission yeast, that incorporates short RNA molecules, such as siRNA and trigger the down-regulation of transcription of a particular gene or genomic region. This is usually accomplished by modification of histone tails, which target the genomic region for heterochromatin formation
AU-rich element	A region in an RNA transcript with frequent A and U nucleotides, such as AUUUA, that targets the RNA for degradation
Piwi-interacting RNA	A small (25–35 nucleotides long) RNA species that is processed from single-stranded precursor RNA, independent of Dicer, and forms a complex with the Piwi protein. piRNAs are probably involved in transposon silencing and stem-cell function

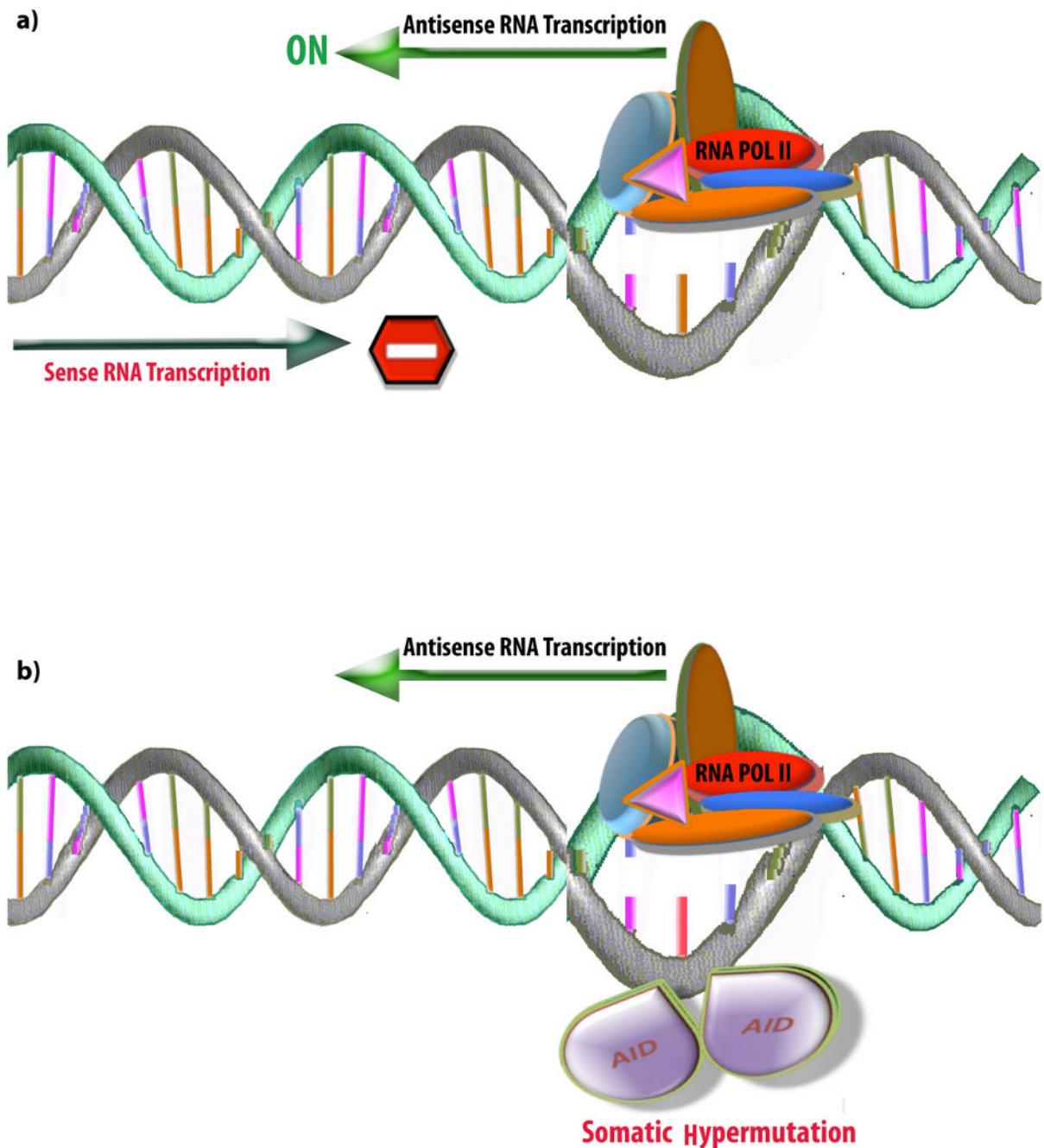
## References

1. Katayama S, et al. Antisense transcription in the mammalian transcriptome. *Science* 2005;309:1564–6. [PubMed: 16141073]
2. Cheng J, et al. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* 2005;308:1149–54. [PubMed: 15790807]
3. Zhang Y, Liu XS, Liu QR, Wei L. Genome-wide in silico identification and analysis of cis natural antisense transcripts (cis-NATs) in ten species. *Nucleic Acids Res* 2006;34:3465–75. [PubMed: 16849434]
4. Rosok O, Sioud M. Systematic identification of sense-antisense transcripts in mammalian cells. *Nat Biotechnol* 2004;22:104–8. [PubMed: 14704709]
5. Ge X, Wu Q, Jung YC, Chen J, Wang SM. A large quantity of novel human antisense transcripts detected by LongSAGE. *Bioinformatics* 2006;22:2475–9. [PubMed: 16895931]

6. Ge X, Rubinstein WS, Jung YC, Wu Q. Genome-wide analysis of antisense transcription with Affymetrix exon array. *BMC Genomics* 2008;9:27. [PubMed: 18211689]
7. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The Antisense Transcriptomes of Human Cells. *Science*. 2008
8. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 2008;322:1845–8. [PubMed: 19056941]
9. Seila AC, et al. Divergent transcription from active promoters. *Science* 2008;322:1849–51. [PubMed: 19056940]
10. Preker P, et al. RNA exosome depletion reveals transcription upstream of active human promoters. *Science* 2008;322:1851–4. [PubMed: 19056938]
11. Finocchiaro G, et al. Localizing hotspots of antisense transcription. *Nucleic Acids Res* 2007;35:1488–500. [PubMed: 17284453]
12. Sun M, Hurst LD, Carmichael GG, Chen J. Evidence for a preferential targeting of 3'-UTRs by cis-encoded natural antisense transcripts. *Nucleic Acids Res* 2005;33:5533–43. [PubMed: 16204454]
13. Okada Y, et al. Comparative expression analysis uncovers novel features of endogenous antisense transcription. *Hum Mol Genet* 2008;17:1631–40. [PubMed: 18283053]
14. Wahlestedt C. Natural antisense and noncoding RNA transcripts as potential drug targets. *Drug Discov Today* 2006;11:503–8. [PubMed: 16713901]
15. Shearwin KE, Callen BP, Egan JB. Transcriptional interference--a crash course. *Trends Genet* 2005;21:339–45. [PubMed: 15922833]
16. Crampton N, Bonass WA, Kirkham J, Rivetti C, Thomson NH. Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. *Nucleic Acids Res* 2006;34:5416–25. [PubMed: 17012275]
17. Prescott EM, Proudfoot NJ. Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci U S A* 2002;99:8796–801. [PubMed: 12077310]
18. Osato N, Suzuki Y, Ieko K, Gojobori T. Transcriptional interferences in cis natural antisense transcripts of humans and mice. *Genetics* 2007;176:1299–306. [PubMed: 17409075]
19. Sastry SS, Hoffman PL. The influence of RNA and DNA template structures during transcript elongation by RNA polymerases. *Biochem Biophys Res Commun* 1995;211:106–14. [PubMed: 7540001]
20. Larijani M, Martin A. Single-stranded DNA structure and positional context of the target cytidine determine the enzymatic efficiency of AID. *Mol Cell Biol* 2007;27:8038–48. [PubMed: 17893327]
21. Perlot T, Li G, Alt FW. Antisense transcripts from immunoglobulin heavy-chain locus V(D)J and switch regions. *Proc Natl Acad Sci U S A* 2008;105:3843–8. [PubMed: 18292225]
22. Ronai D, et al. Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. *J Exp Med* 2007;204:181–90. [PubMed: 17227912]
23. Bolland DJ, et al. Antisense intergenic transcription in V(D)J recombination. *Nat Immunol* 2004;5:630–7. [PubMed: 15107847]
24. Bolland DJ, et al. Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Emu. *Mol Cell Biol* 2007;27:5523–33. [PubMed: 17526723]
25. Roa S, Kuang FL, Scharff MD. Does antisense make sense of AID targeting? *Proc Natl Acad Sci U S A* 2008;105:3661–2. [PubMed: 18322011]
26. Tufarelli C, et al. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet* 2003;34:157–65. [PubMed: 12730694]
27. Yu W, et al. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* 2008;451:202–6. [PubMed: 18185590]
28. Kanduri C. Functional insights into long antisense noncoding RNA Kcnq1ot1 mediated bidirectional silencing. *RNA Biol* 2008;5
29. Ohhata T, Hoki Y, Sasaki H, Sado T. Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. *Development* 2008;135:227–35. [PubMed: 18057104]
30. Bernstein E, Allis CD. RNA meets chromatin. *Genes Dev* 2005;19:1635–55. [PubMed: 16024654]

31. Pandey RR, et al. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* 2008;32:232–46. [PubMed: 18951091]
32. Mohammad F, et al. Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. *Mol Cell Biol* 2008;28:3713–28. [PubMed: 18299392]
33. Imamura T, et al. Non-coding RNA directed DNA demethylation of Sphk1 CpG island. *Biochem Biophys Res Commun* 2004;322:593–600. [PubMed: 15325271]
34. Morris KV, Santoso S, Turner AM, Pastori C, Hawkins PG. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet* 2008;4:e1000258. [PubMed: 19008947]
35. Schwartz JC, et al. Antisense transcripts are targets for activating small RNAs. *Nat Struct Mol Biol* 2008;15:842–8. [PubMed: 18604220]
36. Borel C, et al. Mapping of small RNAs in the human ENCODE regions. *Am J Hum Genet* 2008;82:971–81. [PubMed: 18394580]
37. Kapranov P, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 2007;316:1484–8. [PubMed: 17510325]
38. Goldman SR, Ebright RH, Nickels BE. Direct detection of abortive RNA transcripts in vivo. *Science* 2009;324:927–8. [PubMed: 19443781]
39. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 2002;415:810–3. [PubMed: 11845212]
40. Tomari Y, Zamore PD. Perspective: machines for RNAi. *Genes Dev* 2005;19:517–29. [PubMed: 15741316]
41. Hastings ML, Milcarek C, Martincic K, Peterson ML, Munroe SH. Expression of the thyroid hormone receptor gene, *erbAalpha*, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. *Nucleic Acids Res* 1997;25:4296–300. [PubMed: 9336460]
42. Faghihi MA, et al. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med* 2008;14:723–30. [PubMed: 18587408]
43. Peters NT, Rohrbach JA, Zalewski BA, Byrket CM, Vaughn JC. RNA editing and regulation of *Drosophila* 4f-rnp expression by *sas-10* antisense readthrough mRNA transcripts. *Rna* 2003;9:698–710. [PubMed: 12756328]
44. Ohman M. A-to-I editing challenger or ally to the microRNA process. *Biochimie* 2007;89:1171–6. [PubMed: 17628290]
45. Matsui K, et al. Natural antisense transcript stabilizes inducible nitric oxide synthase messenger RNA in rat hepatocytes. *Hepatology* 2008;47:686–97. [PubMed: 18161049]
46. Uchida T, et al. Prolonged hypoxia differentially regulates hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha expression in lung epithelial cells: implication of natural antisense HIF-1alpha. *J Biol Chem* 2004;279:14871–8. [PubMed: 14744852]
47. Hatzoglou A, et al. Natural antisense RNA inhibits the expression of BCMA, a tumour necrosis factor receptor homologue. *BMC Mol Biol* 2002;3:4. [PubMed: 11960555]
48. Ebralidze AK, et al. PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. *Genes Dev* 2008;22:2085–92. [PubMed: 18676813]
49. Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell* 2005;123:1279–91. [PubMed: 16377568]
50. Zubko E, Meyer P. A natural antisense transcript of the *Petunia hybrida* Sho gene suggests a role for an antisense mechanism in cytokinin regulation. *Plant J* 2007;52:1131–9. [PubMed: 17944812]
51. Katiyar-Agarwal S, Gao S, Vivian-Smith A, Jin H. A novel class of bacteria-induced small RNAs in Arabidopsis. *Genes Dev* 2007;21:3123–34. [PubMed: 18003861]
52. Jin H, Vacic V, Girke T, Lonardi S, Zhu JK. Small RNAs and the regulation of cis-natural antisense transcripts in Arabidopsis. *BMC Mol Biol* 2008;9:6. [PubMed: 18194570]
53. Tam OH, et al. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 2008;453:534–8. [PubMed: 18404147]

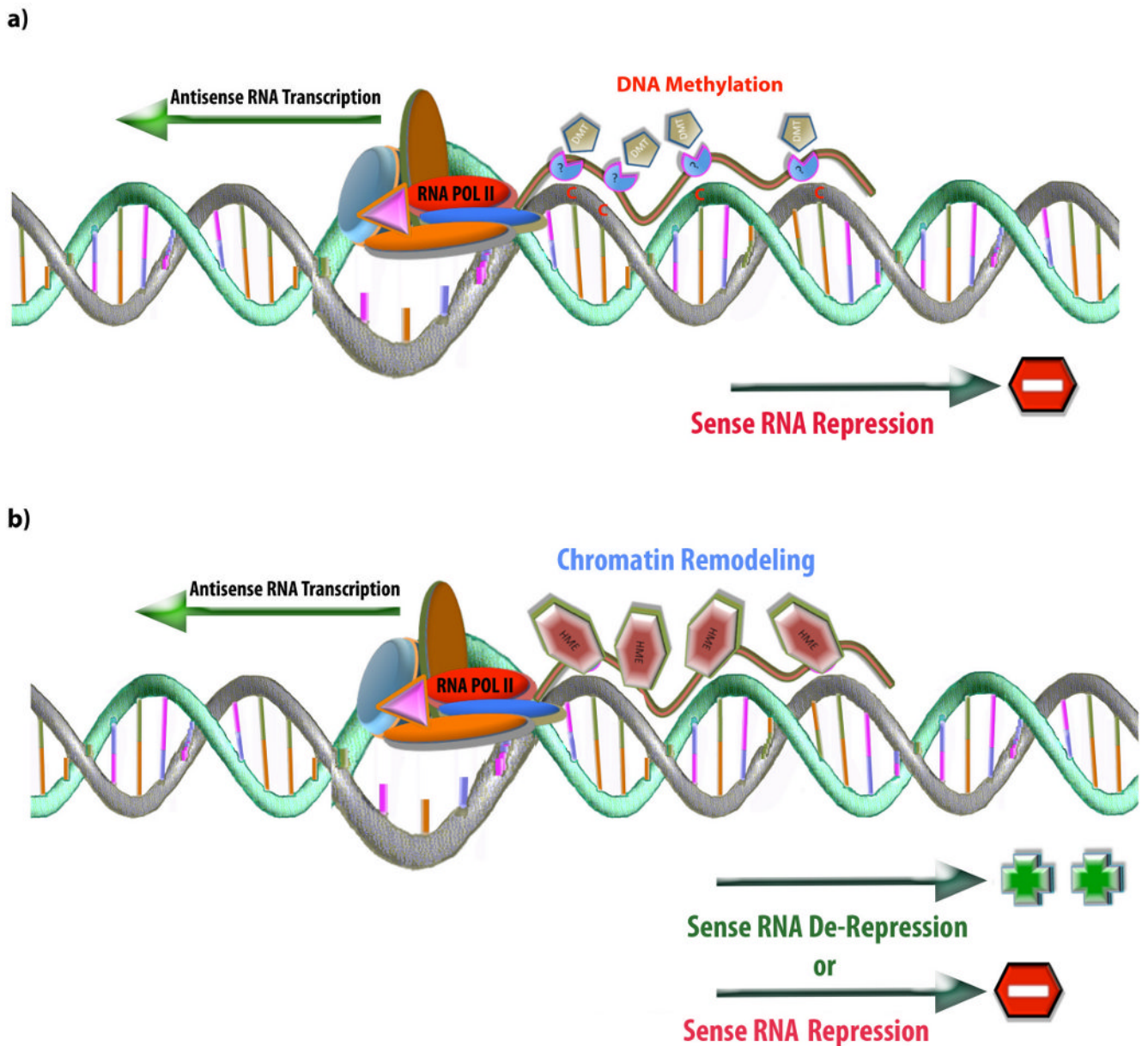
54. Watanabe T, et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 2008;453:539–43. [PubMed: 18404146]
55. Kawaji H, et al. Hidden layers of human small RNAs. *BMC Genomics* 2008;9:157. [PubMed: 18402656]
56. Yang N, Kazazian HH Jr. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* 2006;13:763–71. [PubMed: 16936727]
57. Watanabe T, et al. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev* 2006;20:1732–43. [PubMed: 16766679]
58. Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 2007;318:761–4. [PubMed: 17975059]
59. Klattenhoff C, Theurkauf W. Biogenesis and germline functions of piRNAs. *Development* 2008;135:3–9. [PubMed: 18032451]
60. Okamura K, Balla S, Martin R, Liu N, Lai EC. Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. *Nat Struct Mol Biol* 2008;15:998. [PubMed: 18769470]
61. Czech B, et al. An endogenous small interfering RNA pathway in *Drosophila*. *Nature* 2008;453:798–802. [PubMed: 18463631]
62. Janowski BA, et al. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat Chem Biol* 2007;3:166–73. [PubMed: 17259978]
63. Faghihi MA, Wahlestedt C. RNA interference is not involved in natural antisense mediated regulation of gene expression in mammals. *Genome Biol* 2006;7:R38. [PubMed: 16684369]



**Figure 1. Modulation of transcription**

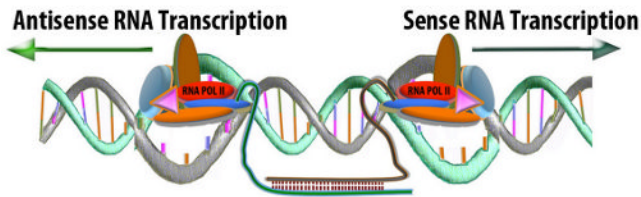
**(a)** Transcription collision model in which RNA synthesis from one DNA strand might clash with transcription from the other strand. According to this model, transcription only occurs in one direction at any given time, and active antisense transcription would suppress sense RNA transcription. Alternatively, supercoiling of DNA strand due to transcription of converging genes is suggested to affect transcription in both directions. **(b)** Genomic DNA rearrangement in hypervariable regions of B and T lymphocytes. Natural antisense transcription from the opposite strand will open the transcription bubble and expose the single-stranded DNA (ssDNA). With the ssDNA exposed, the activation-induced cytidine deaminase (AID) enzyme can bind to DNA and convert deoxycytidine to deoxyuridine, which underlies hypermutation

processes. Considering the small size of the transcription bubble (17 +/- 1 melted bases), which is entirely covered by the elongating RNA polymerase II (RNAPII), it is not clear how AID fits into this region to modify the non-template strand of DNA. Additionally, the selectivity of the somatic hypermutation phenomenon to particular DNA regions suggests that more complex machinery must be involved, in addition to antisense RNA transcription *per se*. Nevertheless, in this proposed model the act of transcription in antisense direction, but not the antisense RNA molecule by itself, would be important for genomic rearrangements.

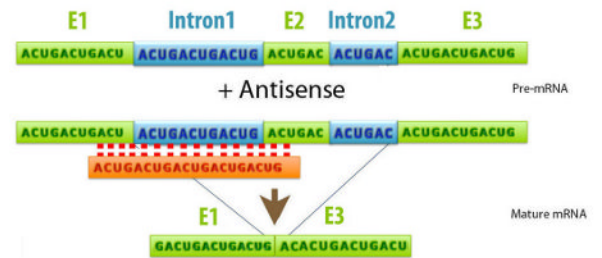


**Figure 2. Induction of chromatin and DNA epigenetic changes by natural antisense transcripts** (a) RNA–DNA interaction model in which newly formed antisense RNA transcript can interact, directly or indirectly, with DNA methyltransferases (DMT) and guide methylation of DNA, resulting in repression of sense RNA transcription (b) Alternatively, newly formed antisense RNA transcript can recruit histone-modifying enzymes (HMEs) and modify chromatin, thereby modulating chromatin architecture and epigenetic memory. Activation of sense RNA transcription, by the means of removal of repressive effects of promoter-directed RNA transcripts, has been observed for several gene loci<sup>34,35,62</sup>. This de-repression effect may be particularly prominent when an antisense RNA is located in or around the transcription start site of the sense gene. Argonaute2 (AGO2) protein is thought to be involved in this largely uncharacterized process (not shown). The resultant chromatin remodelling would repress or de-repress sense RNA transcription. Similar models have been proposed for imprinted genes as well as for mammalian X-chromosome inactivation.

## a) Nuclear RNA-RNA Interactions



## b) Alternative Splicing



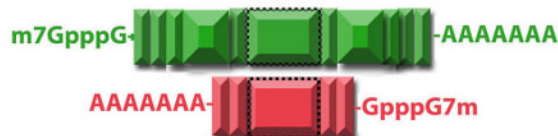
## c) RNA Editing



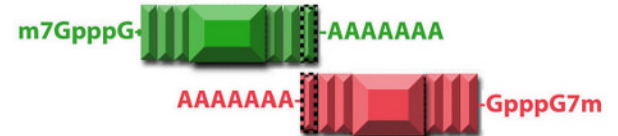
## d) Cytoplasmic RNA-RNA Interactions (Head-to-Head 5' Overlapping)



## e) Cytoplasmic RNA-RNA Interactions (Fully Overlapping)



## f) Cytoplasmic RNA-RNA Interactions (Tail-to-Tail 3' Overlapping)

**Figure-3. Nuclear and cytoplasmic sense-antisense RNA pairing**

(Aa) Nuclear RNA duplex formation might occur locally after transcription, thereby inhibiting sense RNA processing. The RNA processing includes capping, polyadenylation, nuclear localization and transport, which all might be affected by nuclear sense-antisense RNA duplex formation. (Ab) Natural antisense transcripts can cover donor and acceptor splice sites in the sense pre-mRNA transcript to change alternative splicing patterns. (Ac) RNA editing is another possible consequence of nuclear RNA duplex formation, which causes A-to-I editing by the ADAR (adenosine deaminases that act on RNA) enzyme. RNA editing can induce nuclear retention and degradation of hyper-edited transcript by inosine-specific nucleases or alternatively, editing can result in alteration of the amino acid sequence. (B) Cytoplasmic sense-antisense duplex formation can occur in three different formats; (Ba) head-to-head 5' overlapping, (Bb) fully overlapping in which antisense is embedded in the sense transcript or (Bc) tail-to-tail 3' overlapping configuration. Exons are depicted as vertical boxes and the overlapping region is shown with a border. Cytoplasmic sense-antisense RNA duplex formation can possibly exert its own effects on sense mRNA. Hiding or exposing AU-rich elements in sense transcript can affect RNA stability. Changes in the RNA secondary structure,



after binding to antisense RNA, can alter translation, subcellular localization and accessibility of the RNA degradation machinery. According to this model, natural antisense transcripts can potentially 'mask' miRNA-binding sites and release the miRNA-induced block of translation. Finally, Dicer can also process sense–antisense RNA duplexes to endogenous siRNAs.