[Frontiers in Bioscience 17, 300-315, January 1, 2012]

Genome-wide impact of endogenous antisense transcripts in eukaryotes

Koji Numata[1], Hidenori Kiyosawa[2]

[1]Technology and Development Team for Mammalian Cellular Dynamics, BioResource Center (BRC), RIKEN Tsukuba Institute, Ibaraki, 305–0074, Japan, [2]Transdisciplinary Research Integration Center, Research Organization of Information and Systems, National Institute of Genetics, Mishima, Shizuoka, 411–8540, Japan

TABLE OF CONTENTS

1. ABSTRACT

Recent transcriptomic studies revealed that extensive proportions of genomes are transcribed, despite the limited fraction of protein-coding gene loci in the whole genome. Most transcripts are considered to be 'cryptic' output of the genome because of the lack of functional evidence; however, recent progress in molecular analyses has revealed that some of these transcripts at least have functional significance. This review article examines evidence of the functional significance of endogenous *cis*-antisense transcripts, which are the transcriptional output from the opposite strand of annotated genes. These transcripts are one of the most common types of transcripts that do not correspond to any protein-coding loci. Historical molecular studies revealed the existence of antisense transcripts associated with dozens of gene loci, whereas more recent genome-wide studies have shown that many genes have an antisense counterpart thus stimulating investigations into the functional significance of endogenous antisense transcripts. Here, we summarize the recent progress in the genome-wide characterization of the antisense transcriptome, and discuss the biological mechanisms that underlie the regulatory machinery of eukaryotic gene expression with respect to the potential roles of endogenous *cis*-antisense transcripts.

2. INTRODUCTION

The eukaryotic genome produces gene products in a complex regulatory manner. In the last decade, genomic and transcriptomic studies have revealed that eukaryotic gene expression regulation is more complex than previously thought. For example, approximately 90% of human genes are subjected to regulation by alternative splicing (1), which can lead to the generation of a variety of proteins. Diversity in the position of transcription start sites and termination sites within protein-coding gene loci also reflects the complexity of the regulation of gene expression (2). Protein-coding gene loci comprise only a small part of the whole genome, approximately 2–3%, but the actual fraction of the genome contributing to transcriptional output is now known to be much larger. A pilot study targeting 1% of the human genome revealed that more than 90% of the targeted genomic region is transcribed (3). The transcriptional output from outside of the protein-coding gene loci includes the well-characterized structural and functional RNA molecules, such as ribosomal RNA and transfer RNA, as well as the more newly characterized categories of small RNA molecules, including microRNA (miRNA), small-interfering RNA (siRNA), and piwi-interacting RNA (piRNA), which have been shown to play important roles in the regulation of gene expression. These

Table 1. Endogenous antisense transcripts in mammals

Genes	Organisms	References
Dhfr	Mouse	7
Surf-2/Surf-4	Mouse, Human	8,9
c-Myc	Mouse, Human	10
GnRH	Rat	11
c-Myb	Mouse	12
Ribosomal protein L27	Mouse	13
p53	Mouse, Human	14
THRA	Human, Rat	15
ERCC1	Human	16
N-Myc	Mouse, Human	17
eIF2 alpha	Human	18
SC35	Human	19
CD3 zeta/eta/theta	Mouse	20
TYMS	Human	21
BCMA	Human	22
bFGF	Human, Rat	23
WT1	Human	24
Hoxd3	Mouse	25
Hoxa11	Mouse, Human	26
pMCH	Rat	27
Hsp70.2	Mouse	28
Pkn	Mouse	29
MHC	Rat	30
Tgf beta2	Mouse, Rat, Human	31
SMAD5	Human	32
HIF1 alpha	Human	33
UCN	Rat	34
HFE	Human	35
Msx1	Mouse, Rat, Human	36
Raf1	Mouse, Human	37
cTN1	Human, Rat	38
Emx2	Mouse, Human	39
IgH region	Mouse	40
Rfp2/Leu5	Mouse, Human	41

RNA molecules are well-characterized with respect to their functional role in cells as well as the pathways or factors by which they are generated. By comparison, the function of most of the non-protein-coding transcriptome remains uncharacterized. Functionally uncharacterized transcripts are named 'transcripts of unknown function' (TUF) (4), and can be classified into three categories (5). The first category includes transcripts known as antisense transcripts that are complementary to another transcript and can thus form natural double-stranded RNA (dsRNA) molecules. There are two major types of antisense transcripts, *cis*-antisense transcripts (Figure 1A), which are those transcribed from the opposite DNA strand as their target but within the same chromosomal region and form perfect pairs, and *trans*-antisense transcripts (Figure 1B), which are those transcribed from a different location to their targets and generally pair with multiple transcripts with some mismatches (6). The second category of TUF includes non-annotated transcripts that correspond to annotated protein-coding gene loci. They include novel alternative isoforms or transcripts arising with an alternative transcriptional start/end site. The third category includes transcripts mapped to the intronic region of annotated protein-coding genes and those mapped to intergenic regions. This category includes both structural and functional RNA molecules. This review article focuses on the current understanding of endogenous *cis*-antisense transcripts. Recent analyses of the transcriptome in eukaryotes have revealed that endogenous antisense transcription occurs in many regions of the genome and arises from numerous genes and have led to the identification of many *cis*-antisense transcripts associated with specific gene loci.

The presence of antisense transcripts in mammals was first reported by Farnham *et al.* when they found RNA expression from the antisense strand of the 5′ region of *Dhfr*, the gene encoding dihydrofolate reductase (7). This antisense transcript was shown to originate from the protein-coding region in the opposite strand of *Dhfr* according the current genomic annotation. Another report described transcriptional feature of *Surf-2* and *Surf-4* in mouse and suggested that the transcripts from both DNA strands was complementary (8,9). Antisense transcription also arises from the first exon and first intron in the *c-Myc* locus (10). This antisense transcript was shown to not code for protein because it localized to the nucleus and its expression was independent of *c-Myc* expression. Historically, molecular studies revealed the existence of antisense transcripts associated with dozens of gene loci (see Table 1 and references therein). In addition, regulatory roles of antisense transcripts have revealed in various cellular events (Figure 2).

In this review article, we describe the historical background to the genome-wide identification of endogenous antisense transcripts over the last decade and discuss the molecular evidence supporting functional roles of endogenous antisense transcripts in the regulation of eukaryotic gene expression.
11-41)

3. GENOME-WIDE IDENTIFICATION OF ENDOGENOUS ANTISENSE TRANSCRIPTS

3.1. Transcriptome sequencing-based identification of endogenous antisense transcripts

The first computational identification of endogenous antisense transcripts was performed based on an annotated set of mRNAs and a draft human genome sequence (42). To avoid the use of transcripts in the incorrect orientation, the analysis excluded expressed sequence tags (ESTs) and included only those mRNAs annotated with complete CDS (coding sequence) (42). The study identified 87 pairs of sense and antisense transcripts overlapping in *cis*.

Another computer-based study included full-length mRNA sequences and ESTs that had been corrected for transcriptional orientation based on the intronic splice-junction signal (GT and AG) (43). This study identified 144 pairs of sense and antisense transcripts in human and 73 pairs in mouse (43). Some of these pairs have been validated based on confirmation of their expression in cells by RT-PCR (43). Because of incomplete transcriptome data, these studies identified only a small subset of the total number of sense-antisense transcript pairs. Subsequent studies used a more comprehensive dataset of mouse full-length cDNAs and identified ten-fold more *cis*-antisense pairs than previously identified (44,45). By using the full-length cDNA dataset, the study by Kiyosawa *et al.* provided an estimation of the actual number of sense and antisense overlapping transcript pairs in the mouse genome as well as their overlapping patterns relative to the exon/intron structure of the genes; a total of 2481 exon-overlapping pairs that could potentially make natural

double-stranded RNA molecules and 899 non-exon-

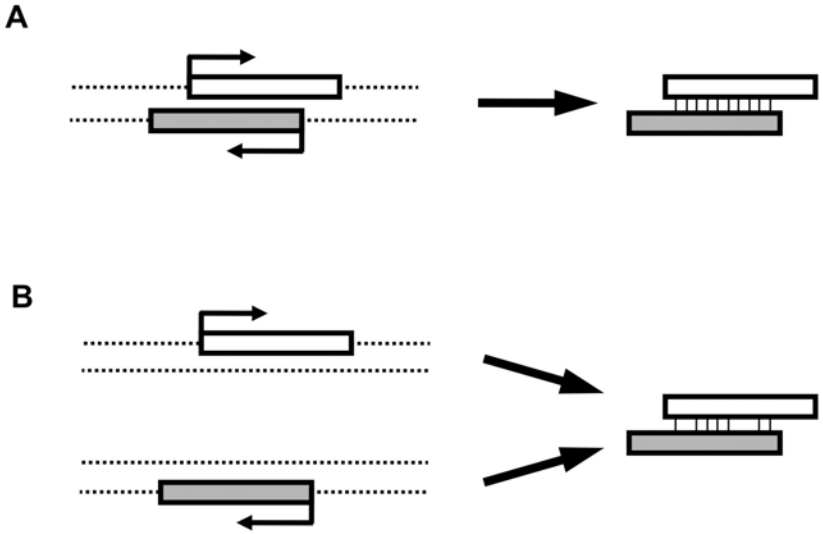

Figure 1. *Cis*- and *trans*-encoded antisense transcript. (A) *Cis*-antisense transcripts. Transcription occurs from the opposite DNA strand as their target but within the same chromosomal region and form perfect pairs. (B) *Trans*-antisense transcripts. Transcription occurs from a different location to their targets and generally pair with multiple transcripts with some mismatches.

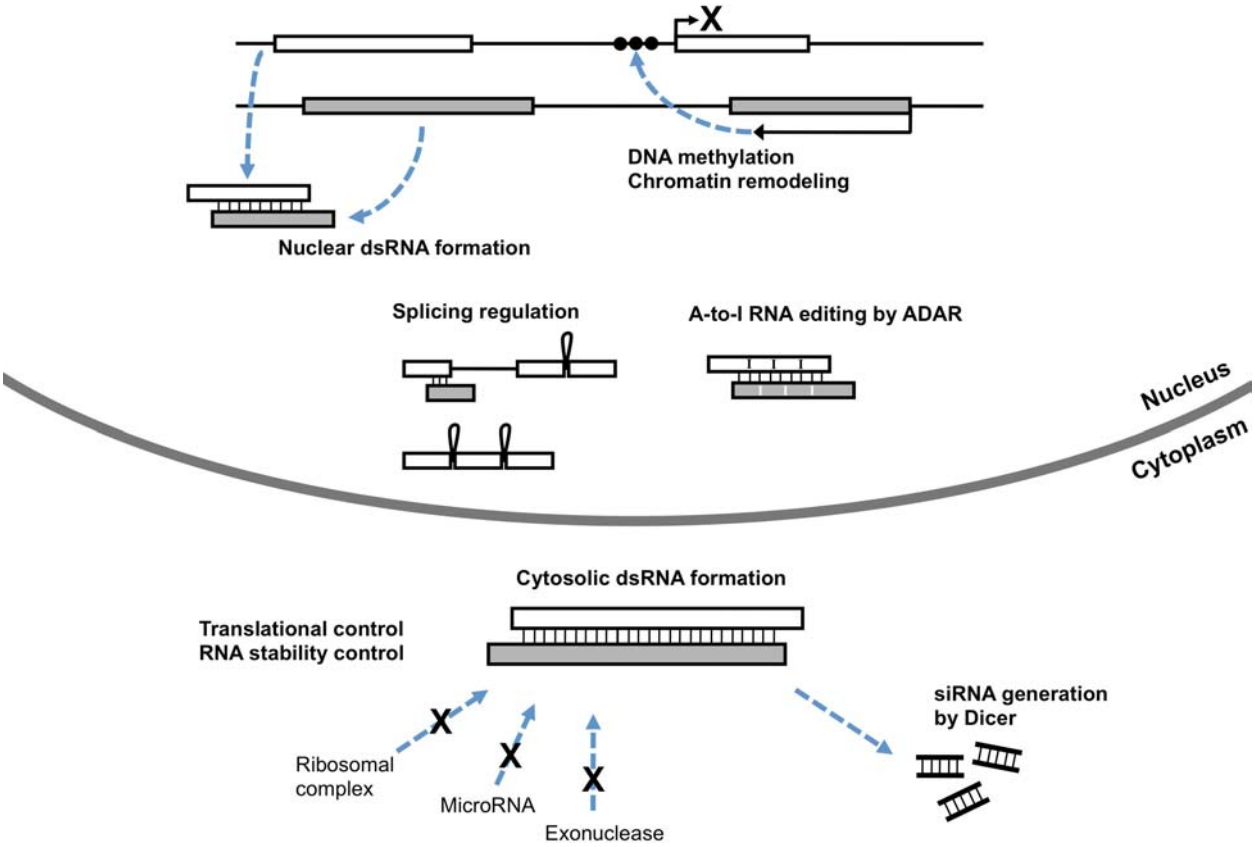

Figure 2. Gene expression regulation mediated by antisense transcripts. There are several antisense-mediated regulatory pathways in cells. Transcription from the opposite strand induces DNA methylation alteration and chromatin remodeling of the surrounding region. Double-stranded RNA (dsRNA) formation in nucleus possibly triggers splicing regulation (by masking target transcript) and RNA editing which converts adenosine to inosine in a site-specific manner mediated by ADAR (adenosine deaminase acting on RNA) family. On the other hand, cytosolic dsRNA formation potentially leads to the translational silencing and RNA stability control by blocking entry of ribosomal complex, microRNA, or exonuclease. It also becomes a template for siRNA generation moderated by RNAi machinery.

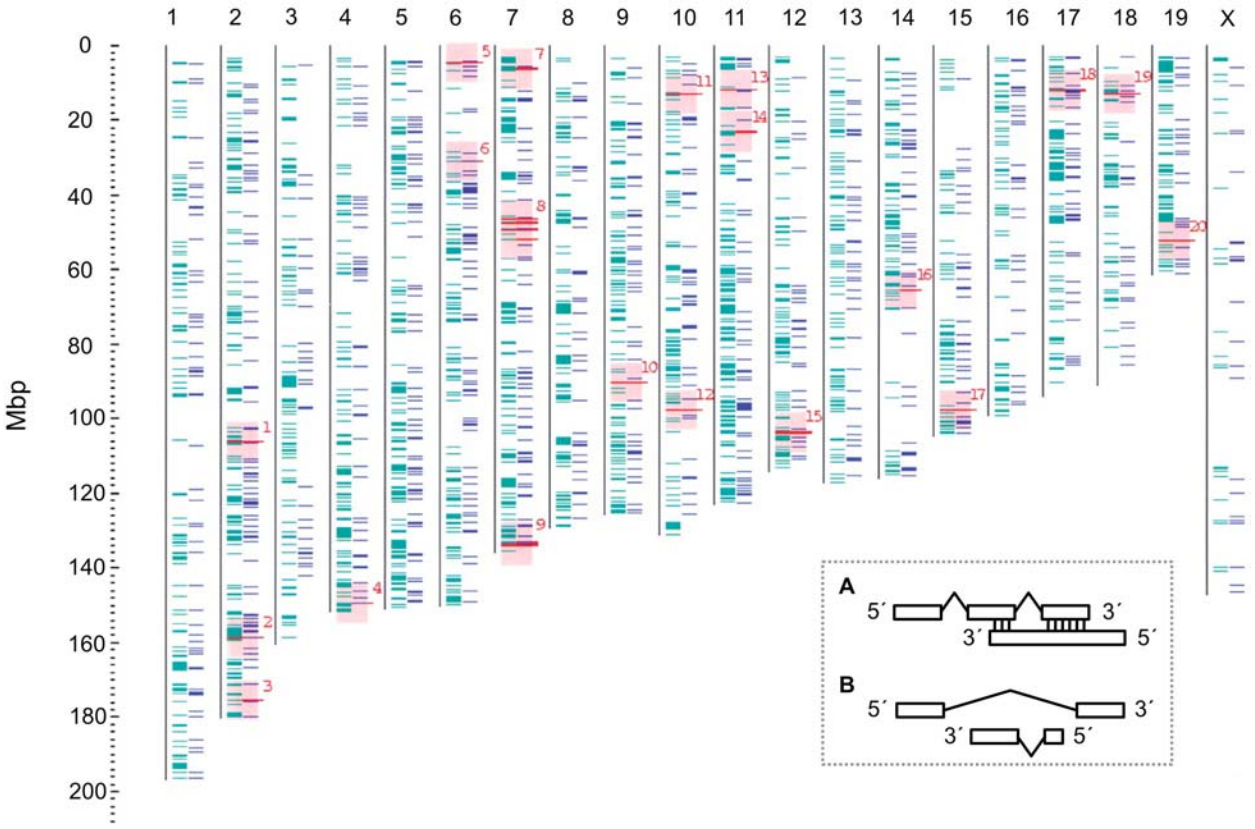

Figure 3. Genome-wide distribution of sense and antisense gene loci in mouse.Location of the pairs of sense and antisense genes on each chromosome. Green lines indicate exon-overlapping pairs that could potentially form natural double-stranded RNA (see example in A), whereas blue lines indicate non-exon-overlapping pairs (see example in B). Known imprinted regions are colored by red with the numbering as follows: Wt1 (1); Nnat (2); Gnas, Gnasxl, Nespas (3); p73 (4); Sgce (5); Copg2, Mit1/lb9 (6); Peg3/Pw1, Usp29, Zim1, Zim3, Zfp264 (7); Frat3, GABRA5, GABRB3, GABRG3, Magel2, Mkrn3/Zfp127, Ndn, Snrpn, Snurf, Ube3a (8); H19, Igf2, Igf2as, Ins2, Ipl/Tssc3, Kvlqt1, Mash2, Nap1l4, Obph1, p57KIP2/Cdkn1c, Slc22a1l, Tapa1/Cd81, Tssc4 (9); Rasgrf1 (10); Zac1 (11); Dcn (12); Meg1/Grb10 (13); U2af1-rs1 region1, U2af1-rs1 region2 (14); Dlk, Meg3/Gtl2 (15); Htr2a (16); Ata3 (17); Igf2r, Mas1, Slc22a2, Slc22a3 (18); Impact (19); Ins1 (20).

overlapping pairs were identified (44). This study also revealed that while endogenous *cis*-antisense transcription is widespread throughout the autosomal regions in mammals, it does not appear to be associated with the X chromosome and is not always enriched within imprinted gene loci (Figure 3) (44). Several studies followed to update the number of the sense-antisense transcript pairs both in humans and mice (46-49). Transcriptome sequencing has revealed the widespread presence of sense-antisense transcript pairs in other model organisms: for example, 1027 pairs in Drosophila (50), 902 in rice (51), and 1340 in Arabidopsis (52). Different approaches to transcriptome sequencing, such as long SAGE (serial analysis of gene expression), have also identified widespread and abundant antisense transcription from gene loci (53,54).

3.2. Microarray-based identification of endogenous antisense transcripts

Microarrays are an efficient tool for transcriptomic studies and can be used to analyze transcript abundance, by using a set of probes each of which is designed to target a specific gene locus. Several studies have used microarrays to detect endogenous antisense transcripts using probes designed to target cDNA (55-59). A custom microarray platform targeting sense and antisense transcripts based on the full-length cDNA collection described above (44), revealed that sense-antisense transcripts are frequently detected in the cDNA pool generated by random primers rather than in the oligo-dT selected cDNA pool (60,61). Together with the finding that some endogenous antisense transcripts localize to the nucleus, the tendency for the sense–antisense transcripts to be poly(A)-negative suggested that a subset of the transcripts in the full-length cDNA collection forms a unique category of RNA molecules distinct from the well-known protein-coding mRNA population. Poly(A)-avoidance of the transcripts is also described in the human transcriptome (as discussed below), as well as in the mouse non-protein-coding transcripts (62). Detailed characteristics of the mammalian transcriptome remain to be elucidated.

Custom microarrays have the potential to detect novel antisense transcripts not within the full-length cDNA collection. A study using a custom microarray approach to target the complementary strand of annotated genes identified novel antisense transcripts arising from cancer-related genes that had an altered pattern of expression in tumor tissue (63). This approach also identified antisense transcripts that had been detected by EST analyses in the human transcriptome, but not in the mouse transcriptome. These antisense transcripts might be functionally conserved between human and mouse, but are unable to be identified solely on the basis of cDNA and EST analyses.

The microarray approach for transcriptome studies usually uses one probe (or one set of probes) designed for each gene locus, whereas the tiling microarray approach uses probes to achieve more widespread coverage of genomic regions, such as whole chromosomes, and can detect at high-resolution the transcriptional activity of the target region in a strand-specific manner (64-66). The benefit of the tiling array approach is its ability to detect transcriptional activity in any part of the genome without being dependent on the completeness of the transcriptome data, and genomic annotation of computationally predicted genes. Tiling array expression profiling studies targeting non-mammalian model species, including yeast, Arabidopsis, and rice, revealed that a significant proportion of their genomes are transcribed both from sense and antisense strands (67-69). In humans, chromosome-wide tiling array analyses showed that transcription, generating both polyadenylated and non-polyadenylated transcripts, was widespread throughout the genome (64,70). These studies also showed that more than half of the transcripts are located in the nucleus, suggesting that they are most likely non-protein-coding transcripts (64,70). Another important finding was that more than half of all human genes are expressed from the opposite strand.

In summary, microarray analysis has been used to determine the relative expression levels of endogenous antisense transcripts but because of limited probe coverage their transcription start and stop sites have not been identified. One approach to address this issue is transcriptome sequencing, either by full-length cDNA Sanger sequencing or the recently developed strand-specific deep sequencing (71,72).

4. ANTISENSE TRANSCRIPTS AS EPIGENETIC REGULATORS

The best known role of endogenous antisense transcripts is the control of gene expression by transcriptional repression. In eukaryotes, the regulation of gene expression is controlled by transcription factors as well as by epigenetic changes in the surrounding structure either by covalent histone modifications or by DNA cytosine methylation. Endogenous antisense transcription sometimes influences the epigenetic remodeling of the neighboring genomic region (73). This section introduces typical examples of antisense transcripts that act as epigenetic regulators.

4.1. X-chromosome inactivation

The most studied example of endogenous antisense transcripts, as well as long non-protein-coding RNA, as epigenetic regulators are those arising from the *Xist–Tsix* locus, a key regulator of mammalian X-chromosome dosage compensation. In mammals, to avoid the dosage differences of X-linked genes between male and female, one of the two X-chromosomes in female cells is transcriptionally inactivated. This process, called X-inactivation, occurs through X chromosome heterochromatin formation, which is initiated by *Xist*, a long non-protein-coding RNA initially identified as an inactive X-specific transcript (74,75). *Xist* is encoded on the X chromosome and is a 17 kb transcript in humans and a 15 kb transcript in mouse (76,77). *Xist* is upregulated at the blastocyst stage of early embyogenesis and inactivates the X chromosome in *cis* by spreading over the entire chromosome (78). Because cells lacking *Xist* do not exhibit X-inactivation, and because expression of *Xist* from an autosomal region inactivates the expression of surrounding genes (79), *Xist* expression is considered to be sufficient to achieve X-inactivation. *Xist* expression, and hence the transcriptional activity of the X chromosome, is controlled by the expression of *Tsix*, the antisense counterpart of *Xist*. *Xist* expression is repressed by *Tsix* on the future active X chromosome, whereas on the future inactive X chromosome *Tsix* expression is repressed thereby allowing *Xist* to be expressed and initiate X inactivation. The *Tsix*-mediated repression of *Xist* expression is accomplished by modulation of chromatin structure through covalent histone modification within the neighboring genomic region, especially within the promoter region of *Xist*. Mutation of *Tsix* causes changes in the methylation status of histone H3 lysine 9 (H3K9), H3 lysine 4 (H3K4), H3 lysine 27 (H3K27), and also of DNA cytosine methylation of the *Xist* region (80-83). The premature termination of *Tsix* transcription across the *Xist* promoter leads to activation of *Xist* expression and loss of X inactivation thus indicating that the *Tsix* is crucial to X-inactivation (84).

4.2. Antisense transcripts in imprinted loci

Genomic imprinting is the phenomenon whereby one of the two alleles of the diploid genome are expressed in a parental origin-specific manner (85). Because differences in DNA methylation status between two alleles are often observed in imprinted regions, the regulation of imprinted gene expression is thought to occur via epigenetic molecular machinery. In the mouse genome, approximately 90 imprinted gene loci have been identified thus far (http://www.har.mrc.ac.uk/research/genomic_imprinting/).

Endogenous antisense transcription in commonly observed in imprinted gene loci; 10–15% of them are associated with antisense transcripts (85). At present, 11 of those in human, and 13 of those in mouse are listed in the ncRNAimprint database, a comprehensive database of mammalian imprinted noncoding RNAs (86). Functional analyses of antisense transcripts in imprinted regions have been carried out in depth for sense-antisense pairs, especially for the *Igf2r–Air* and *Kcnq1–Kcnq1ot1* sense-antisense pairs.

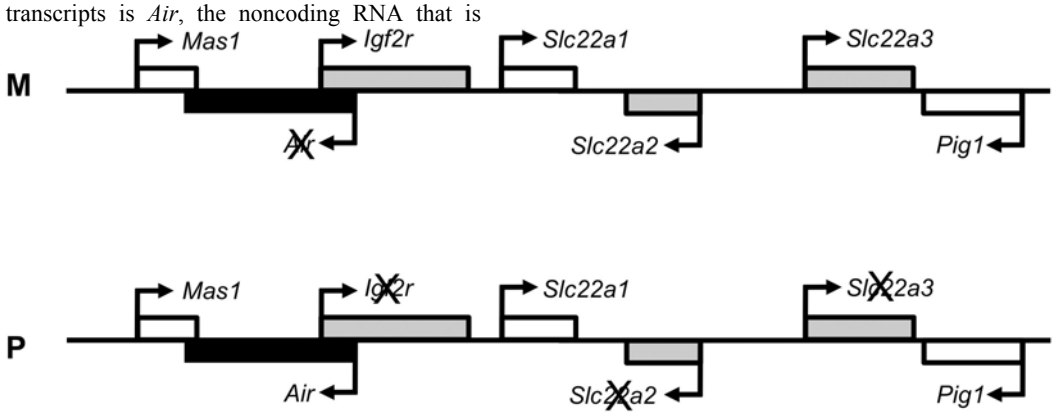

Figure 4. Genomic structure of the region controlled by *Air*. The region controlled by *Air* is illustrated (chromosome 17 in mouse). Expression of *Air* (black bar) represses three genes (*Igf2r*, *Slc22a2*, and *Slc22a3*; gray-colored) on the paternally (P) inherited chromosome. These three genes are expressed only from the maternal (M) allele, where Air transcription is repressed. White-colored genes are expressed in a bi-allelic manner.

One of the most intensely studied imprinted antisense transcripts is *Air*, the noncoding RNA that is antisense to the insulin-like growth factor 2 receptor (*Igf2r*) gene. In mouse, *Igf2r* is maternally expressed in most tissues, but biallelically expressed only in brain (87). *Air* is transcribed from the opposite strand of the intronic region of *Igf2r*, and consists of a 100 kb unspliced transcript (88-90). This widespread transcription occurs only from the paternal allele, suggesting that *Air* represses paternal *Igf2r* expression in *cis*. Because premature truncation of the *Air* transcript, such as promoter deletion or polyadenylation cassette insertion, leads to upregulation of the paternal expression of *Igf2r/Slc22a2/Slc22a3* (88,91,92), *Air* probably represses the paternal alleles of those genes in *cis* (Figure 4). Details of the molecular mechanism underlying the *Air*-mediated repression of *Igf2r*, and the nearby imprinted genes *Slc22a2* and *Slc22a3,* remain unclear. However, alterations to the chromatin structure are most likely involved because *Air* accumulates at the promoter of *Slc22a3* and co-localizes with the methylation of H3K9 (93).

Similar antisense-mediated regulation is also reported for *Kcnq1* (potassium voltage-gated channel, subfamily Q, member 1) and its antisense partner, *Kcnq1ot1* (*Kcnq1* overlapping transcript). *Kcnq1* exhibits maternal-specific expression, whereas *Kcnq1ot1* is expressed only from the paternal allele. Transcription of *Kcnq1ot1* arises from the opposite strand of the intronic region of *Kcnq1*, and suppresses several genes within a 780 kb span (94). Truncation of *Kcnq1ot*1 results in alteration of the allele-specific expression of adjacent genes (95-97). At the promoter region of these genes, the trimethylation of both H3K27 and H3K9 is enriched (98,99), suggesting that *Kcnq1ot1* modifies the chromatin structure of the target genes. Moreover, *Kcnq1ot1* interacts with methyltransferase G9a, which methylates histone H3, as well as with the PRC2 complex (100), suggesting that this antisense transcript acts as a suppressor of gene expression *in cis*.

The ubiquitin protein ligase E3A (*Ube3a*) gene antisense transcript, *Ube3a-ATS* (also known as *LNCAT*) (101,102), is another well-known example of an antisense transcript in an imprinted loci; however, the chromatin structure and the effect of truncation of this antisense transcript are not known. *Ube3a* is the causative gene in Angelman syndrome, and is maternally expressed only in brain, both in mouse and human (103,104). The tissue-specific allele-specific expression of *Ube3a* is concomitant with its brain-specific antisense transcript, started from the neighboring *Snrpn* locus, which is the region responsible for Prader-Willi syndrome (101,102). Transcription of the opposite DNA strand spans approximately 1000 kb, and the region contains repeat elements and individual snoRNA (small nucleolar RNA)-encoding units (102,105,106) (Figure 5), suggesting that allele-specific expression of this region might be regulated by a very complex mechanism.

In mammals, mono-allelic gene expression can occur not only in a parent-of-origin specific manner but also in random (107). The recently developed high-throughput transcriptome sequencing approach has identified many gene loci exhibiting parent-of-origin allelic expression (108) as well as strain-biased allelic imbalance of expression (109). Further experimental analyses are needed to understand the relationships between mono-allelic gene expression, antisense transcription, and epigenetic modification.

4.3. Antisense transcripts associated with epigenetic alteration in disease

Antisense transcript also affects epigenetic alterations in diseases. For example, in alpha-thalassemia patients transcriptional read-through of the human LUC7-derived antisense transcript induces methylation of the promoter of the hemoglobin alpha 2 (HBA2) gene, which abolishes HBA2 expression (110). In leukemia, epigenetic silencing of the tumor suppressor gene *p15* (cyclin-dependent kinase inhibitor 2B) is caused by the upregulation of its antisense RNA couterpart, which leads to heterochromatin formation within the p15 gene and a subsequent decrease in its expression (111). Since antisense transcripts have also been identified in other human disease gene loci encoding transcription factors, including *p21*

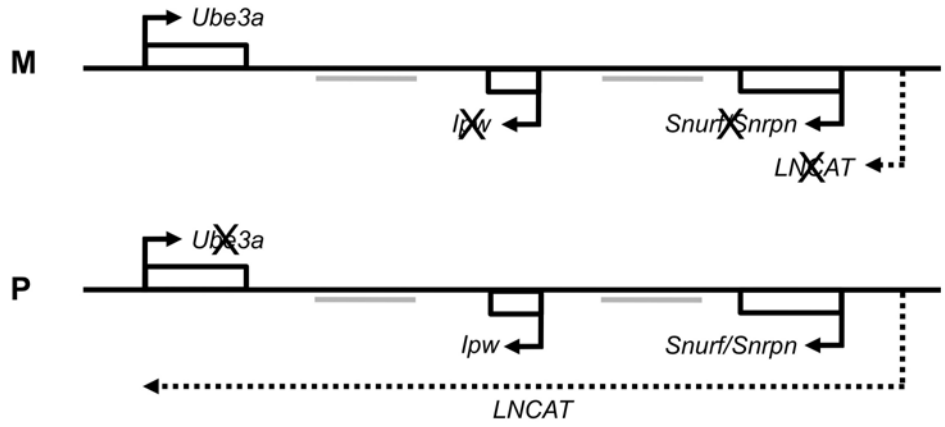

Figure 5. Antisense transcript of *Ube3a* and the surrounding region. *Ube3a* is maternally expressed in the brain. This allele-specific expression is associated with the antisense transcript, the transcription of which is a part of widespread transcription of this region, named LNCAT (dashed line), starting upstream of the *Snurf/Snrpn* gene locus on the paternally inherited chromosome. In most tissues, except brain, *Ube3a* is bi-allelically expressed in the absence of LNCAT expression. *Snurf/Snrpn* is expressed paternally in most tissues. Grey lines indicate the region of the snoRNA (small nucleolar RNA)-encoded cluster.

(112), *PU.1* (113), *p53* (14), *c-Myc* (10), and *HIF* (33), expression analysis of both strands of gene loci is crucial in investigations of disease transcriptomes. Indeed, microarray and strand-specific deep sequencing have identified several gene loci in which sense and antisense expression is reversely correlated between normal and tumor conditions (63,72).

5. MEDIATION OF SMALL RNA FORMATION BY ANTISENSE TRANSCRIPTION

The simultaneous transcription of both DNA strands can lead to the formation of double-stranded RNA (dsRNA). Major pathways involving dsRNA are the RNA interference (RNAi) pathway and the RNA editing pathway (114). Several RNAi-related events associated with endogenous antisense transcript have been recently identified. In this section, we describe how *cis*-encoded antisense transcript triggers the formation of endogenous siRNAs. We also describe a new class of short RNA molecules that frequently occur near the boundary region of protein-coding gene loci and whose function was previously unknown, until now.

5.1. Endogenous siRNA formation in plants

In plants, it is well known that endogenous siRNA, which is capable of dsRNA formation, is generated by transcription from both DNA strands. For example, in Arabidopsis, two types of siRNA (21 nt and 24 nt) are generated from a pair of natural *cis*-antisense transcripts arising from the overlapping region between 1-pyrroline-5-carboxylate dehydrogenase (*P5CDH*), a stress-related gene, and *SRO5*, a gene of unknown function, which is annotated as a stress-related gene. Expression of *SRO5* is induced by salt stress, and triggers siRNA production, which promotes the degradation of *P5CDH* mRNA. The siRNA-induced reduction in the level of P5CDH mRNA leads to a decrease in proline degradation and thus an increase in the salt stress tolerance of Arabidopsis (115). Two endogenous siRNAs, derived from natural *cis*-antisense transcripts from the reversely overlapping gene pairs ATGB2/PPRL and SRRLK/AtRAP, are reported in plants with bacterial infections (116,117). This finding suggests that endogenous siRNAs are induced in response to the stress of bacterial infection. Formation of these siRNAs involves the dicer-like (DCL) ribonuclease type III enzyme (75,116,117).

Moreover, several genome-wide studies revealed that small RNAs are often located within the overlapping region of two genes (118), suggesting that there might be other examples of siRNAs derived by endogenous antisense transcription. One study estimated that more than half of sense-antisense transcript loci are capable of forming siRNAs, suggesting that the formation of siRNAs from natural *cis*-antisense transcripts is a common mechanism of gene expression regulation in plants.

5.2. Endogenous siRNA formation in animals

The formation of endogenous siRNAs from *cis*-encoded sense-antisense transcripts is also observed in animals. In Drosophila, 21-nt-long siRNAs are derived from a region enriched with overlapping sense and antisense transcripts (119). High-throughput sequencing of the mouse oocyte small RNA also revealed the existence of two classes of endogenous small RNAs: that is, piRNA and siRNA (120,121). piRNAs are associated with repeat sequences and act as retrotransposon regulators in germline development (122), whereas siRNAs, also found in germ cells, are generated from the natural *cis*-antisense transcripts, which are naturally occurring dsRNA molecules (120). This study found 17 loci in which endogenous siRNAs could potentially be processed from the natural dsRNA molecules generated by *cis*-encoded antisense transcripts; for example, from the overlapping *Pdzd11* and *Kif4* gene pair on the X chromosome (88). The level of these siRNAs is decreased with the concomitant loss of Dicer activity, and the loss of Dicer activity also leads to the upregulation of the host genes, suggesting that these siRNAs are generated by a cytoplasmic Dicer-

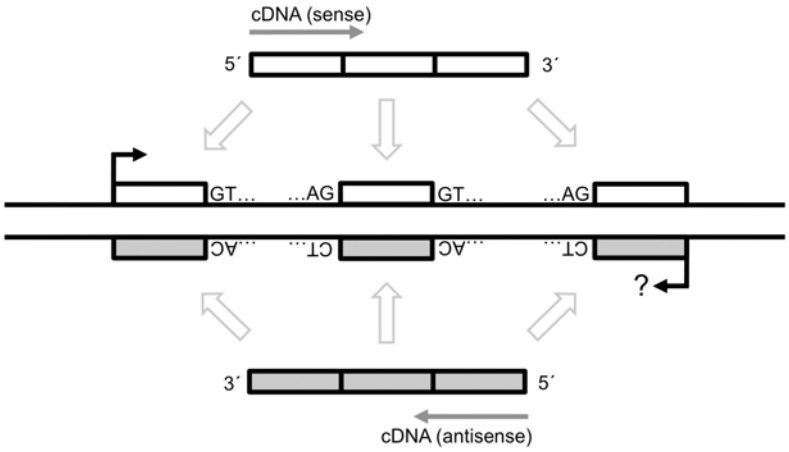

Figure 6. Schematic illustration of sense and antisense transcripts sharing the same splice junction. In several gene loci, cDNA derived from the antisense strand has the same exon/intron structure as that of the sense gene. This type of antisense transcript might be spliced in the same region as the sense transcript at the splicing signal CT/AC or it might be newly synthesized from the mRNA template of the sense strand using an as yet unknown RdRp-like RNA copying mechanism (see main text).

dependent pathway. Cawlile *et al.* also identified endogenous siRNA generated from the *Slc34a1* (Na/phosphate cotransporter) locus in mouse kidney and testis (123). Intriguingly, the orientation of siRNAs differs among tissues, suggesting that different regulatory pathways for the generation of siRNAs exist in different tissues. Endogenous siRNAs are generated not only from *cis*-antisense transcript loci, but also from expressed pseudogenes and inverted repeats that are capable of forming natural dsRNA molecules (120). A Dicer-dependent small RNA is also detected in the *Xist–Tsix* region in murine ES cells; this study revealed that an endogenous siRNA represses *Xist* expression from the active X chromosome suggesting that this siRNA is generated through *cis*-antisense transcription and mediates the epigenetic modifications regulating *Xist* expression (124).

Taken together, these molecular findings support the existence of endogenous siRNAs associated with sense and antisense transcript pairs, and suggest that a subset of the large amount of *cis*-encoded antisense transcripts that have been identified but have no known function play a role in the generation of regulatory siRNA molecules.

5.3. Short RNA molecules of unknown function

Biochemically well-characterized small RNAs (siRNA, miRNA, and piRNA) are usually processed into 20–40 nt long molecules. However, recent genome-scale studies revealed that a distinct population of the short RNA molecules is likely to be expressed from the genome, although the molecular association with endogenous antisense transcript remains to be elucidated.

In humans, tiling arrays and deep sequencing targeting the short RNA fraction (≤200 nt) revealed that short RNAs of various sizes (22 to 200 nt) are likely to be expressed specifically from the promoter and terminator region of protein-coding genes. These short RNAs are named promoter-associated short RNAs (PASR) and termini associated short RNAs (TASR) (70,125). Deep sequencing of the murine ES cell transcriptome detected a similar size range of short RNAs that are enriched near the boundary region of the protein-coding genes, which comprise more than half of the total genes of the genome (126). These short RNAs are transcribed from both DNA strands and co-localize with RNA polymerase II and H3K4 trimethylation, suggesting that they are expressed from transcriptionally active regions. Lack of Dicer activity does not lead to the loss of these short RNAs, suggesting that they are not generated through an RNAi-dependent pathway.

Although their biological role and the pathway by which these short RNAs are generated is still uncharacterized, it is possible that they are associated with strand-specific transcription of the surrounding region because Northern blot analyses of target regions of the sense-antisense transcript pairs detected a similar size range of short RNAs (60,61). Some were detected only in the cytosolic RNA fraction, suggesting that the short RNAs have diverse biological roles.

6. EXPLANATION OF UNEXPECTED ENDOGENOUS ANTISENSE TRANSCRIPTS IN MAMMALS

Although a subset of the endogenous antisense transcriptome appears to be unprocessed, like *Air*, the genomic alignment pattern suggests that many of these transcripts are likely to be spliced (44). These spliced antisense transcripts are thought to arise from discrete protein-coding gene loci on the opposite strand of other gene loci. However, several loci show an intriguing pattern in which the exon/intron structure of both strands is identical (22,34,38,127). In such cases, the donor and accepter sites for splicing of the antisense transcript are reciprocally complementary to that of the sense gene (GT/AG in the sense and CT/AC in antisense) (Figure 6).

There are two possible mechanisms underlying the generation of these antisense transcripts. First, the antisense transcript could be spliced in the same region as the sense transcript through reorganization of unusual splice site signals (CT/AC) by the spliceosomal complex. Second, the antisense transcript could be newly synthesized from the mRNA template of the sense strand. The latter mechanism premises the presence of RNA-dependent RNA polymerase (RdRp) activity, and such activity has not been reported in mammals, except for in the telomeric region (128).

In the beta-globin locus, antisense transcripts are still stable, even though treatment with actinomycin D inhibits *de novo* transcription (127). In addition, the antisense transcript to the Troponin I gene possesses a poly(U)-stretch in its 5′ region, suggesting that it might be generated from the poly(A)-tail of Troponin I (129). Taken together, these findings suggest that some antisense transcripts might be generated using the sense mRNA counterpart as a template.

Recently, massively parallel, single-molecule sequencing has revealed that a subset of short RNA molecules (≤ 200 nt) have a poly(U) tail at their 5′ end (130). These short RNAs tend to be enriched within the opposite strand of the 3′ end of annotated gene loci, and thus are likely to be generated by RdRp-like activity that reverse transcribes using the 3′ poly(A)-tail of mRNA as a template. Although the biological significance of 5′ poly(U) RNA molecules remains unclear, it is possible that they act as a primer, or that they are simply an incomplete product of an unknown RNA copying mechanisms that generates antisense transcripts from the sense mRNA template. These observations suggest that endogenous antisense transcripts are generated through similar RNA-copying mechanisms in humans.

7. CONCLUSION

This article reviews current knowledge of eukaryotic endogenous antisense transcripts including their identification by recent genome-wide computer-based studies and biological roles revealed by analyses of individual loci. We focus on two categories of functional antisense transcripts, those which induce epigenetic changes and endogenous small RNA molecules. Both categories of antisense transcripts have a role in repression of gene expression. A recent study has revealed another category of antisense transcripts that affects the stability of the sense mRNA counterpart (131,132). This finding suggests that the antisense transcriptome in eukaryotic cells contains antisense transcripts with a repressive role but also another category with a regulatory role yet to be elucidated. Because several aspects of the mechanisms underlying the function of antisense transcripts are unclear, for example poly(A)-avoidance and a possible RNA-copying mechanism, a variety of approached are required (genetic, biochemical, and genomic) to fully understand the functions of the eukaryotic transcriptome, especially that of the endogenous antisense transcripts and non-protein-coding transcripts.

8. REFERENCES

1. Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B.: Alternative isoform regulation in human tissue transcriptomes. *Nature*, 456, 470-476 (2008)

2. Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impiombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminiecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavesi, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.: The transcriptional landscape of the mammalian genome. *Science*, 309, 1559-1563 (2005)

3. Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., Kuehn, M.S., Taylor, C.M., Neph, S., Koch, C.M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J.A., Andrews, R.M., Flicek, P., Boyle, P.J., Cao, H., Carter, N.P., Clelland, G.K., Davis, S.,

Day, N., Dhami, P., Dillon, S.C., Dorschner, M.O., Fiegler, H., Giresi, P.G., Goldy, J., Hawrylycz, M., Haydock, A., Humbert, R., James, K.D., Johnson, B.E., Johnson, E.M., Frum, T.T., Rosenzweig, E.R., Karnani, N., Lee, K., Lefebvre, G.C., Navas, P.A., Neri, F., Parker, S.C., Sabo, P.J., Sandstrom, R., Shafer, A., Vetrie, D., Weaver, M., Wilcox, S., Yu, M., Collins, F.S., Dekker, J., Lieb, J.D., Tullius, T.D., Crawford, G.E., Sunyaev, S., Noble, W.S., Dunham, I., Denoeud, F., Reymond, A., Kapranov, P., Rozowsky, J., Zheng, D., Castelo, R., Frankish, A., Harrow, J., Ghosh, S., Sandelin, A., Hofacker, I.L., Baertsch, R., Keefe, D., Dike, S., Cheng, J., Hirsch, H.A., Sekinger, E.A., Lagarde, J., Abril, J.F., Shahab, A., Flamm, C., Fried, C., Hackermuller, J., Hertel, J., Lindemeyer, M., Missal, K., Tanzer, A., Washietl, S., Korbel, J., Emanuelsson, O., Pedersen, J.S., Holroyd, N., Taylor, R., Swarbreck, D., Matthews, N., Dickson, M.C., Thomas, D.J., Weirauch, M.T., Gilbert, J., Drenkow, J., Bell, I., Zhao, X., Srinivasan, K.G., Sung, W.K., Ooi, H.S., Chiu, K.P., Foissac, S., Alioto, T., Brent, M., Pachter, L., Tress, M.L., Valencia, A., Choo, S.W., Choo, C.Y., Ucla, C., Manzano, C., Wyss, C., Cheung, E., Clark, T.G., Brown, J.B., Ganesh, M., Patel, S., Tammana, H., Chrast, J., Henrichsen, C.N., Kai, C., Kawai, J., Nagalakshmi, U., Wu, J., Lian, Z., Lian, J., Newburger, P., Zhang, X., Bickel, P., Mattick, J.S., Carninci, P., Hayashizaki, Y., Weissman, S., Hubbard, T., Myers, R.M., Rogers, J., Stadler, P.F., Lowe, T.M., Wei, C.L., Ruan, Y., Struhl, K., Gerstein, M., Antonarakis, S.E., Fu, Y., Green, E.D., Karaoz, U., Siepel, A., Taylor, J., Liefer, L.A., Wetterstrand, K.A., Good, P.J., Feingold, E.A., Guyer, M.S., Cooper, G.M., Asimenos, G., Dewey, C.N., Hou, M., Nikolaev, S., Montoya-Burgos, J.I., Loytynoja, A., Whelan, S., Pardi, F., Massingham, T., Huang, H., Zhang, N.R., Holmes, I., Mullikin, J.C., Ureta-Vidal, A., Paten, B., Seringhaus, M., Church, D., Rosenbloom, K., Kent, W.J., Stone, E.A., Batzoglou, S., Goldman, N., Hardison, R.C., Haussler, D., Miller, W., Sidow, A., Trinklein, N.D., Zhang, Z.D., Barrera, L., Stuart, R., King, D.C., Ameur, A., Enroth, S., Bieda, M.C., Kim, J., Bhinge, A.A., Jiang, N., Liu, J., Yao, F., Vega, V.B., Lee, C.W., Ng, P., Yang, A., Moqtaderi, Z., Zhu, Z., Xu, X., Squazzo, S., Oberley, M.J., Inman, D., Singer, M.A., Richmond, T.A., Munn, K.J., Rada-Iglesias, A., Wallerman, O., Komorowski, J., Fowler, J.C., Couttet, P., Bruce, A.W., Dovey, O.M., Ellis, P.D., Langford, C.F., Nix, D.A., Euskirchen, G., Hartman, S., Urban, A.E., Kraus, P., Van Calcar, S., Heintzman, N., Kim, T.H., Wang, K., Qu, C., Hon, G., Luna, R., Glass, C.K., Rosenfeld, M.G., Aldred, S.F., Cooper, S.J., Halees, A., Lin, J.M., Shulha, H.P., Xu, M., Haidar, J.N., Yu, Y., Iyer, V.R., Green, R.D., Wadelius, C., Farnham, P.J., Ren, B., Harte, R.A., Hinrichs, A.S., Trumbower, H., Clawson, H., Hillman-Jackson, J., Zweig, A.S., Smith, K., Thakkapallayil, A., Barber, G., Kuhn, R.M., Karolchik, D., Armengol, L., Bird, C.P., de Bakker, P.I., Kern, A.D., Lopez-Bigas, N., Martin, J.D., Stranger, B.E., Woodroffe, A., Davydov, E., Dimas, A., Eyras, E., Hallgrimsdottir, I.B., Huppert, J., Zody, M.C., Abecasis, G.R., Estivill, X., Bouffard, G.G., Guan, X., Hansen, N.F., Idol, J.R., Maduro, V.V., Maskeri, B., McDowell, J.C., Park, M., Thomas, P.J., Young, A.C., Blakesley, R.W., Muzny, D.M., Sodergren, E., Wheeler, D.A., Worley, K.C., Jiang, H., Weinstock, G.M., Gibbs, R.A., Graves, T., Fulton, R., Mardis, E.R., Wilson, R.K., Clamp, M., Cuff, J., Gnerre, S., Jaffe, D.B., Chang, J.L., Lindblad-Toh, K., Lander, E.S., Koriabine, M., Nefedov, M., Osoegawa, K., Yoshinaga, Y., Zhu, B. and de Jong, P.J.: Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447, 799-816 (2007)

4. Willingham, A.T. and Gingeras, T.R.: TUF love for "junk" DNA. *Cell*, 125, 1215-1220 (2006)

5. Gingeras, T.R.: Origin of phenotypes: genes and transcripts. *Genome Res*, 17, 682-690 (2007)

6. Vanhee-Brossollet, C. and Vaquero, C.: Do natural antisense transcripts make sense in eukaryotes? *Gene*, 211, 1-9 (1998)

7. Farnham, P.J., Abrams, J.M. and Schimke, R.T.: Opposite-strand RNAs from the 5' flanking region of the mouse dihydrofolate reductase gene. *Proc Natl Acad Sci U S A*, 82, 3978-3982 (1985)

8. Williams, T. and Fried, M.: A mouse locus at which transcription from both DNA strands produces mRNAs complementary at their 3' ends. *Nature*, 322, 275-279 (1986)

9. Williams, T., Yon, J., Huxley, C. and Fried, M.: The mouse surfeit locus contains a very tight cluster of four "housekeeping" genes that is conserved through evolution. *Proc Natl Acad Sci U S A*, 85, 3527-3530 (1988)

10. Nepveu, A. and Marcu, K.B.: Intragenic pausing and anti-sense transcription within the murine c-myc locus. *EMBO J*, 5, 2859-2865 (1986)

11. Adelman, J.P., Bond, C.T., Douglass, J. and Herbert, E.: Two mammalian genes transcribed from opposite strands of the same DNA locus. *Science*, 235, 1514-1517 (1987)

12. Bender, T.P., Thompson, C.B. and Kuehl, W.M.: Differential expression of c-myb mRNA in murine B lymphomas by a block to transcription elongation. *Science*, 237, 1473-1476 (1987)

13. Belhumeur, P., Lussier, M. and Skup, D.: Expression of naturally occurring RNA molecules complementary to the murine L27' ribosomal protein mRNA. *Gene*, 72, 277-285 (1988)

14. Khochbin, S. and Lawrence, J.J.: An antisense RNA involved in p53 mRNA maturation in murine erythroleukemia cells induced to differentiate. *EMBO J*, 8, 4107-4114 (1989)

15. Lazar, M.A., Hodin, R.A., Darling, D.S. and Chin, W.W.: A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat c-erbA alpha transcriptional unit. *Mol Cell Biol*, 9, 1128-1136 (1989)

16. van Duin, M., van Den Tol, J., Hoeijmakers, J.H., Bootsma, D., Rupp, I.P., Reynolds, P., Prakash, L. and Prakash, S.: Conserved pattern of antisense overlapping transcription in the homologous human ERCC-1 and yeast RAD10 DNA repair gene regions. *Mol Cell Biol*, 9, 1794-1798 (1989)

17. Krystal, G.W., Armstrong, B.C. and Battey, J.F.: N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol Cell Biol*, 10, 4180-4191 (1990)

18. Silverman, T.A., Noguchi, M. and Safer, B.: Role of sequences within the first intron in the regulation of expression of eukaryotic initiation factor 2 alpha. *J Biol Chem*, 267, 9738-9742 (1992)

19. Fu, X.D. and Maniatis, T.: Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science*, 256, 535-538 (1992)

20. Lerner, A., D'Adamio, L., Diener, A.C., Clayton, L.K. and Reinherz, E.L.: CD3 zeta/eta/theta locus is colinear with and transcribed antisense to the gene encoding the transcription factor Oct-1. *J Immunol*, 151, 3152-3162 (1993)

21. Dolnick, B.J.: Cloning and characterization of a naturally occurring antisense RNA to human thymidylate synthase mRNA. *Nucleic Acids Res*, 21, 1747-1752 (1993)

22. Laabi, Y., Gras, M.P., Brouet, J.C., Berger, R., Larsen, C.J. and Tsapis, A.: The BCMA gene, preferentially expressed during B lymphoid maturation, is bidirectionally transcribed. *Nucleic Acids Res*, 22, 1147-1154 (1994)

23. Murphy, P.R. and Knee, R.S.: Identification and characterization of an antisense RNA transcript (gfg) from the human basic fibroblast growth factor gene. *Mol Endocrinol*, 8, 852-859 (1994)

24. Eccles, M.R., Grubb, G., Ogawa, O., Szeto, J. and Reeve, A.E.: Cloning of novel Wilms tumor gene (WT1) cDNAs; evidence for antisense transcription of WT1. *Oncogene*, 9, 2059-2063 (1994)

25. Bedford, M., Arman, E., Orr-Urtreger, A. and Lonai, P.: Analysis of the Hoxd-3 gene: structure and localization of its sense and natural antisense transcripts. *DNA Cell Biol*, 14, 295-304 (1995)

26. Hsieh-Li, H.M., Witte, D.P., Weinstein, M., Branford, W., Li, H., Small, K. and Potter, S.S.: Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility. *Development*, 121, 1373-1385 (1995)

27. Hervieu, G. and Nahon, J.L.: Pro-melanin concentrating hormone messenger ribonucleic acid and peptides expression in peripheral tissues of the rat. *Neuroendocrinology*, 61, 348-364 (1995)

28. Murashov, A.K. and Wolgemuth, D.J.: Distinct transcripts are recognized by sense and antisense riboprobes for a member of the murine HSP70 gene family, HSP70.2, in various reproductive tissues. *Mol Reprod Dev*, 43, 17-24 (1996)

29. Murashov, A.K. and Wolgemuth, D.J.: Sense and antisense transcripts of the developmentally regulated murine hsp70.2 gene are expressed in distinct and only partially overlapping areas in the adult brain. *Brain Res Mol Brain Res*, 37, 85-95 (1996)

30. Luther, H.P., Haase, H., Hohaus, A., Beckmann, G., Reich, J. and Morano, I.: Characterization of naturally occurring myosin heavy chain antisense mRNA in rat heart. *J Cell Biochem*, 70, 110-120 (1998)

31. Coker, R.K., Laurent, G.J., Dabbagh, K., Dawson, J. and McAnulty, R.J.: A novel transforming growth factor beta2 antisense transcript in mammalian lung. *Biochem J*, 332 (Pt 2), 297-301 (1998)

32. Zavadil, J., Svoboda, P., Liang, H., Kottickal, L.V. and Nagarajan, L.: An antisense transcript to SMAD5 expressed in fetal and tumor tissues. *Biochem Biophys Res Commun*, 255, 668-672 (1999)

33. Thrash-Bingham, C.A. and Tartof, K.D.: aHIF: a natural antisense transcript overexpressed in human renal cancer and during hypoxia. *J Natl Cancer Inst*, 91, 143-151 (1999)

34. Shi, M., Yan, X., Ryan, D.H. and Harris, R.B.: Identification of urocortin mRNA antisense transcripts in rat tissue. *Brain Res Bull*, 53, 317-324 (2000)

35. Thenie, A.C., Gicquel, I.M., Hardy, S., Ferran, H., Fergelot, P., Le Gall, J.Y. and Mosser, J.: Identification of an endogenous RNA transcribed from the antisense strand of the HFE gene. *Hum Mol Genet*, 10, 1859-1866 (2001)

36. Blin-Wakkach, C., Lezot, F., Ghoul-Mazgar, S., Hotton, D., Monteiro, S., Teillaud, C., Pibouin, L., Orestes-Cardoso, S., Papagerakis, P., Macdougall, M., Robert, B. and Berdal, A.: Endogenous Msx1 antisense transcript: *in vivo* and *in vitro* evidences, structure, and potential involvement in skeleton development in mammals. *Proc Natl Acad Sci U S A*, 98, 7336-7341 (2001)

37. Gray, T.A., Azama, K., Whitmore, K., Min, A., Abe, S. and Nicholls, R.D.: Phylogenetic conservation of the makorin-2 gene, encoding a multiple zinc-finger protein, antisense to the RAF1 proto-oncogene. *Genomics*, 77, 119-126 (2001)

38. Podlowski, S., Bramlage, P., Baumann, G., Morano, I. and Luther, H.P.: Cardiac troponin I sense-antisense RNA duplexes in the myocardium. *J Cell Biochem*, 85, 198-207 (2002)

39. Noonan, F.C., Goodfellow, P.J., Staloch, L.J., Mutch, D.G. and Simon, T.C.: Antisense transcripts at the EMX2 locus in human and mouse. *Genomics*, 81, 58-66 (2003)

40. Bolland, D.J., Wood, A.L., Johnston, C.M., Bunting, S.F., Morgan, G., Chakalova, L., Fraser, P.J. and Corcoran,

A.E.: Antisense intergenic transcription in V(D)J recombination. *Nat Immunol*, 5, 630-637 (2004)

41. Corcoran, M.M., Hammarsund, M., Zhu, C., Lerner, M., Kapanadze, B., Wilson, B., Larsson, C., Forsberg, L., Ibbotson, R.E., Einhorn, S., Oscier, D.G., Grander, D. and Sangfelt, O.: DLEU2 encodes an antisense RNA for the putative bicistronic RFP2/LEU5 gene in humans and mouse. *Genes Chromosomes Cancer*, 40, 285-297 (2004)

42. Lehner, B., Williams, G., Campbell, R.D. and Sanderson, C.M.: Antisense transcripts in the human genome. *Trends Genet*, 18, 63-65 (2002)

43. Shendure, J. and Church, G.M.: Computational discovery of sense-antisense transcription in the human and mouse genomes. *Genome Biol*, 3, RESEARCH0044 (2002)

44. Kiyosawa, H., Yamanaka, I., Osato, N., Kondo, S. and Hayashizaki, Y.: Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Res*, 13, 1324-1334 (2003)

45. Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.: Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, 420, 563-573 (2002)

46. Engstrom, P.G., Suzuki, H., Ninomiya, N., Akalin, A., Sessa, L., Lavorgna, G., Brozzi, A., Luzi, L., Tan, S.L., Yang, L., Kunarso, G., Ng, E.L., Batalov, S., Wahlestedt, C., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., Wells, C., Bajic, V.B., Orlando, V., Reid, J.F., Lenhard, B. and Lipovich, L.: Complex Loci in human and mouse genomes. *PLoS Genet*, 2, e47 (2006)

47. Sun, M., Hurst, L.D., Carmichael, G.G. and Chen, J.: Evidence for a preferential targeting of 3'-UTRs by cis-encoded natural antisense transcripts. *Nucleic Acids Res*, 33, 5533-5543 (2005)

48. Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.: Antisense transcription in the mammalian transcriptome. *Science*, 309, 1564-1566 (2005)

49. Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z. and Rowley, J.D.: Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res*, 32, 4812-4820 (2004)

50. Misra, S., Crosby, M.A., Mungall, C.J., Matthews, B.B., Campbell, K.S., Hradecky, P., Huang, Y., Kaminker, J.S., Millburn, G.H., Prochnik, S.E., Smith, C.D., Tupy, J.L., Whitfied, E.J., Bayraktaroglu, L., Berman, B.P., Bettencourt, B.R., Celniker, S.E., de Grey, A.D., Drysdale, R.A., Harris, N.L., Richter, J., Russo, S., Schroeder, A.J., Shu, S.Q., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W.M., Rubin, G.M. and Lewis, S.E.: Annotation of the Drosophila melanogaster euchromatic genome: a systematic review. *Genome Biol*, 3, RESEARCH0083 (2002)

51. Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.J., Ohtsuki, K., Shishiki, T., Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., Hayashizaki, Y. and Yasunishi, A.: Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science*, 301, 376-379 (2003)

52. Wang, X.J., Gaasterland, T. and Chua, N.H.: Genome-wide prediction and identification of cis-natural antisense transcripts in Arabidopsis thaliana. *Genome Biol*, 6, R30 (2005)

53. Ge, X., Wu, Q., Jung, Y.C., Chen, J. and Wang, S.M.: A large quantity of novel human antisense transcripts detected by LongSAGE. *Bioinformatics*, 22, 2475-2479 (2006)

54. Wahl, M.B., Heinzmann, U. and Imai, K.: LongSAGE analysis revealed the presence of a large number of novel antisense genes in the mouse genome. *Bioinformatics*, 21, 1389-1392 (2005)

55. Ge, X., Rubinstein, W.S., Jung, Y.C. and Wu, Q.: Genome-wide analysis of antisense transcription with Affymetrix exon array. *BMC Genomics*, 9, 27 (2008)

56. Gyorffy, A., Surowiak, P., Tulassay, Z. and Gyorffy, B.: Highly expressed genes are associated with inverse antisense transcription in mouse. *J Genet*, 86, 103-109 (2007)

57. Vallon-Christersson, J., Staaf, J., Kvist, A., Medstrand, P., Borg, A. and Rovira, C.: Non-coding antisense transcription detected by conventional and single-stranded cDNA microarray. *BMC Genomics*, 8, 295 (2007)

58. Oeder, S., Mages, J., Flicek, P. and Lang, R.: Uncovering information on expression of natural antisense transcripts in Affymetrix MOE430 datasets. *BMC Genomics*, 8, 200 (2007)

59. Werner, A., Schmutzler, G., Carlile, M., Miles, C.G. and Peters, H.: Expression profiling of antisense transcripts on DNA arrays. *Physiol Genomics*, 28, 294-300 (2007)

60. Okada, Y., Tashiro, C., Numata, K., Watanabe, K., Nakaoka, H., Yamamoto, N., Okubo, K., Ikeda, R., Saito, R., Kanai, A., Abe, K., Tomita, M. and Kiyosawa, H.: Comparative expression analysis uncovers novel features of endogenous antisense transcription. *Hum Mol Genet*, 17, 1631-1640 (2008)

61. Kiyosawa, H., Mise, N., Iwase, S., Hayashizaki, Y. and Abe, K.: Disclosing hidden transcripts: mouse natural sense-antisense transcripts tend to be poly(A) negative and nuclear localized. *Genome Res*, 15, 463-474 (2005)

62. Furuno, M., Pang, K.C., Ninomiya, N., Fukuda, S., Frith, M.C., Bult, C., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., Mattick, J.S. and Suzuki, H.: Clusters of internally primed transcripts reveal novel long noncoding RNAs. *PLoS Genet*, 2, e37 (2006)

63. Numata, K., Osada, Y., Okada, Y., Saito, R., Hiraiwa, N., Nakaoka, H., Yamamoto, N., Watanabe, K., Okubo, K., Kohama, C., Kanai, A., Abe, K. and Kiyosawa, H.: Identification of novel endogenous antisense transcripts by DNA microarray analysis targeting complementary strand of annotated genes. *BMC Genomics*, 10, 392 (2009)

64. Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G., Sementchenko, V., Piccolboni, A., Bekiranov, S., Bailey, D.K., Ganesh, M., Ghosh, S., Bell, I., Gerhard, D.S. and Gingeras, T.R.: Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science*, 308, 1149-1154 (2005)

65. Kampa, D., Cheng, J., Kapranov, P., Yamanaka, M., Brubaker, S., Cawley, S., Drenkow, J., Piccolboni, A., Bekiranov, S., Helt, G., Tammana, H. and Gingeras, T.R.: Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Res*, 14, 331-342 (2004)

66. Kapranov, P., Cawley, S.E., Drenkow, J., Bekiranov, S., Strausberg, R.L., Fodor, S.P. and Gingeras, T.R.: Large-scale transcriptional activity in chromosomes 21 and 22. *Science*, 296, 916-919 (2002)

67. David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C.J., Bofkin, L., Jones, T., Davis, R.W. and Steinmetz, L.M.: A high-resolution map of transcription in the yeast genome. *Proc Natl Acad Sci U S A*, 103, 5320-5325 (2006)

68. Li, L., Wang, X., Stolc, V., Li, X., Zhang, D., Su, N., Tongprasit, W., Li, S., Cheng, Z., Wang, J. and Deng, X.W.: Genome-wide transcription analyses in rice using tiling microarrays. *Nat Genet*, 38, 124-129 (2006)

69. Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S.X., Lam, B., Sakano, H., Wu, T., Yu, G., Miranda, M., Quach, H.L., Tripp, M., Chang, C.H., Lee, J.M., Toriumi, M., Chan, M.M., Tang, C.C., Onodera, C.S., Deng, J.M., Akiyama, K., Ansari, Y., Arakawa, T., Banh, J., Banno, F., Bowser, L., Brooks, S., Carninci, P., Chao, Q., Choy, N., Enju, A., Goldsmith, A.D., Gurjal, M., Hansen, N.F., Hayashizaki, Y., Johnson-Hopson, C., Hsuan, V.W., Iida, K., Karnes, M., Khan, S., Koesema, E., Ishida, J., Jiang, P.X., Jones, T., Kawai, J., Kamiya, A., Meyers, C., Nakajima, M., Narusaka, M., Seki, M., Sakurai, T., Satou, M., Tamse, R., Vaysberg, M., Wallender, E.K., Wong, C., Yamamura, Y., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.: Empirical analysis of transcriptional activity in the Arabidopsis genome. *Science*, 302, 842-846 (2003)

70. Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermuller, J., Hofacker, I.L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., Sementchenko, V., Tammana, H. and Gingeras, T.R.: RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*, 316, 1484-1488 (2007)

71. Levin, J.Z., Yassour, M., Adiconis, X., Nusbaum, C., Thompson, D.A., Friedman, N., Gnirke, A. and Regev, A.: Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat Methods*, 7, 709-715 (2010)

72. Morrissy, A.S., Morin, R.D., Delaney, A., Zeng, T., McDonald, H., Jones, S., Zhao, Y., Hirst, M. and Marra,

M.A.: Next-generation tag sequencing for cancer gene expression profiling. *Genome Res*, 19, 1825-1835 (2009)

73. Su, W.Y., Xiong, H. and Fang, J.Y.: Natural antisense transcripts regulate gene expression in an epigenetic manner. *Biochem Biophys Res Commun*, 396, 177-181 (2010)

74. Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D. and Rastan, S.: Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature*, 351, 329-331 (1991)

75. Borsani, G., Tonlorenzi, R., Simmler, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., Willard, H.F., Avner, P. and Ballabio, A.: Characterization of a murine gene expressed from the inactive X chromosome. *Nature*, 351, 325-329 (1991)

76. Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J. and Willard, H.F.: The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, 71, 527-542 (1992)

77. Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S. and Rastan, S.: The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, 71, 515-526 (1992)

78. Heard, E. and Disteche, C.M.: Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev*, 20, 1848-1867 (2006)

79. Wutz, A. and Jaenisch, R.: A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell*, 5, 695-705 (2000)

80. Shibata, S. and Yokota, T.: Alteration of histone tail modifications in the Xist locus in wild-type and Tsix-mutant male embryonic stem cells during differentiation. *Exp Anim*, 57, 153-157 (2008)

81. Navarro, P., Page, D.R., Avner, P. and Rougeulle, C.: Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. *Genes Dev*, 20, 2787-2792 (2006)

82. Sun, B.K., Deaton, A.M. and Lee, J.T.: A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol Cell*, 21, 617-628 (2006)

83. Sado, T., Hoki, Y. and Sasaki, H.: Tsix silences Xist through modification of chromatin structure. *Dev Cell*, 9, 159-165 (2005)

84. Ohhata, T., Hoki, Y., Sasaki, H. and Sado, T.: Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. *Development*, 135, 227-235 (2008)

85. Reik, W. and Walter, J.: Genomic imprinting: parental influence on the genome. *Nat Rev Genet*, 2, 21-32 (2001)

86. Zhang, Y., Guan, D.G., Yang, J.H., Shao, P., Zhou, H. and Qu, L.H.: ncRNAimprint: a comprehensive database of mammalian imprinted noncoding RNAs. *RNA*, 16, 1889-1901 (2010)

87. Barlow, D.P., Stoger, R., Herrmann, B.G., Saito, K. and Schweifer, N.: The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature*, 349, 84-87 (1991)

88. Sleutels, F., Zwart, R. and Barlow, D.P.: The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*, 415, 810-813 (2002)

89. Lyle, R., Watanabe, D., te Vruchte, D., Lerchner, W., Smrzka, O.W., Wutz, A., Schageman, J., Hahner, L., Davies, C. and Barlow, D.P.: The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat Genet*, 25, 19-21 (2000)

90. Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F. and Barlow, D.P.: Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature*, 389, 745-749 (1997)

91. Zwart, R., Sleutels, F., Wutz, A., Schinkel, A.H. and Barlow, D.P.: Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. *Genes Dev*, 15, 2361-2366 (2001)

92. Wutz, A., Theussl, H.C., Dausman, J., Jaenisch, R., Barlow, D.P. and Wagner, E.F.: Non-imprinted Igf2r expression decreases growth and rescues the Tme mutation in mice. *Development*, 128, 1881-1887 (2001)

93. Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R. and Fraser, P.: The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science*, 322, 1717-1720 (2008)

94. Smilinich, N.J., Day, C.D., Fitzpatrick, G.V., Caldwell, G.M., Lossie, A.C., Cooper, P.R., Smallwood, A.C., Joyce, J.A., Schofield, P.N., Reik, W., Nicholls, R.D., Weksberg, R., Driscoll, D.J., Maher, E.R., Shows, T.B. and Higgins, M.J.: A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proc Natl Acad Sci U S A*, 96, 8064-8069 (1999)

95. Shin, J.Y., Fitzpatrick, G.V. and Higgins, M.J.: Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. *EMBO J*, 27, 168-178 (2008)

96. Mancini-Dinardo, D., Steele, S.J., Levorse, J.M., Ingram, R.S. and Tilghman, S.M.: Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev*, 20, 1268-1282 (2006)

97. Fitzpatrick, G.V., Soloway, P.D. and Higgins, M.J.: Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat Genet*, 32, 426-431 (2002)

98. Terranova, R., Yokobayashi, S., Stadler, M.B., Otte, A.P., van Lohuizen, M., Orkin, S.H. and Peters, A.H.: Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev Cell*, 15, 668-679 (2008)

99. Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y. and Feil, R.: Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet*, 36, 1296-1300 (2004)

100. Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-Dinardo, D. and Kanduri, C.: Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell*, 32, 232-246 (2008)

101. Landers, M., Bancescu, D.L., Le Meur, E., Rougeulle, C., Glatt-Deeley, H., Brannan, C., Muscatelli, F. and Lalande, M.: Regulation of the large (approximately 1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn. *Nucleic Acids Res*, 32, 3480-3492 (2004)

102. Le Meur, E., Watrin, F., Landers, M., Sturny, R., Lalande, M. and Muscatelli, F.: Dynamic developmental regulation of the large non-coding RNA associated with the mouse 7C imprinted chromosomal region. *Dev Biol*, 286, 587-600 (2005)

103. Rougeulle, C., Glatt, H. and Lalande, M.: The Angelman syndrome candidate gene, UBE3A/E6-AP, is imprinted in brain. *Nat Genet*, 17, 14-15 (1997)

104. Vu, T.H. and Hoffman, A.R.: Imprinting of the Angelman syndrome gene, UBE3A, is restricted to brain. *Nat Genet*, 17, 12-13 (1997)

105. Vitali, P., Royo, H., Marty, V., Bortolin-Cavaille, M.L. and Cavaille, J.: Long nuclear-retained non-coding RNAs and allele-specific higher-order chromatin organization at imprinted snoRNA gene arrays. *J Cell Sci*, 123, 70-83 (2010)

106. Landers, M., Calciano, M.A., Colosi, D., Glatt-Deeley, H., Wagstaff, J. and Lalande, M.: Maternal disruption of Ube3a leads to increased expression of Ube3a-ATS in trans. *Nucleic Acids Res*, 33, 3976-3984 (2005)

107. Krueger, C. and Morison, I.M.: Random monoallelic expression: making a choice. *Trends Genet*, 24, 257-259 (2008)

108. Gregg, C., Zhang, J., Weissbourd, B., Luo, S., Schroth, G.P., Haig, D. and Dulac, C.: High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science*, 329, 643-648 (2010)

109. Babak, T., Deveale, B., Armour, C., Raymond, C., Cleary, M.A., van der Kooy, D., Johnson, J.M. and Lim, L.P.: Global survey of genomic imprinting by transcriptome sequencing. *Curr Biol*, 18, 1735-1741 (2008)

110. Tufarelli, C., Stanley, J.A., Garrick, D., Sharpe, J.A., Ayyub, H., Wood, W.G. and Higgs, D.R.: Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet*, 34, 157-165 (2003)

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

112. Morris, K.V., Santoso, S., Turner, A.M., Pastori, C. and Hawkins, P.G.: Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet*, 4, e1000258 (2008)

113. Ebralidze, A.K., Guibal, F.C., Steidl, U., Zhang, P., Lee, S., Bartholdy, B., Jorda, M.A., Petkova, V., Rosenbauer, F., Huang, G., Dayaram, T., Klupp, J., O'Brien, K.B., Will, B., Hoogenkamp, M., Borden, K.L., Bonifer, C. and Tenen, D.G.: PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. *Genes Dev*, 22, 2085-2092 (2008)

114. Lavorgna, G., Dahary, D., Lehner, B., Sorek, R., Sanderson, C.M. and Casari, G.: In search of antisense. *Trends Biochem Sci*, 29, 88-94 (2004)

115. Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R. and Zhu, J.K.: Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*, 123, 1279-1291 (2005)

116. Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A. and Jin, H.: A novel class of bacteria-induced small RNAs in Arabidopsis. *Genes Dev*, 21, 3123-3134 (2007)

117. Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., Borsani, O., Villegas, A., Jr., Zhu, J.K., Staskawicz, B.J. and Jin, H.: A pathogen-inducible endogenous siRNA in plant immunity. *Proc Natl Acad Sci U S A*, 103, 18002-18007 (2006)

118. Jin, H., Vacic, V., Girke, T., Lonardi, S. and Zhu, J.K.: Small RNAs and the regulation of cis-natural antisense transcripts in Arabidopsis. *BMC Mol Biol*, 9, 6 (2008)

119. Okamura, K., Balla, S., Martin, R., Liu, N. and Lai, E.C.: Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. *Nat Struct Mol Biol*, 15, 581-590 (2008)

120. Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M.A., Sakaki, Y. and Sasaki, H.: Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, 453, 539-543 (2008)

121. Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N. and Imai, H.: 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*, 20, 1732-1743 (2006)

122. Klattenhoff, C. and Theurkauf, W.: Biogenesis and germline functions of piRNAs. *Development*, 135, 3-9 (2008)

123. Carlile, M., Swan, D., Jackson, K., Preston-Fayers, K., Ballester, B., Flicek, P. and Werner, A.: Strand selective generation of endo-siRNAs from the Na/phosphate transporter gene Slc34a1 in murine tissues. *Nucleic Acids Res*, 37, 2274-2282 (2009)

124. Ogawa, Y., Sun, B.K. and Lee, J.T.: Intersection of the RNA interference and X-inactivation pathways. *Science*, 320, 1336-1341 (2008)

125. Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project: Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature*, 457, 1028-1032 (2009)

126. Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A. and Sharp, P.A.: Divergent transcription from active promoters. *Science*, 322, 1849-1851 (2008)

127. Volloch, V.: Cytoplasmic synthesis of globin RNA in differentiated murine erythroleukemia cells: possible involvement of RNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A*, 83, 1208-1212 (1986)

128. Maida, Y., Yasukawa, M., Furuuchi, M., Lassmann, T., Possemato, R., Okamoto, N., Kasim, V., Hayashizaki, Y., Hahn, W.C. and Masutomi, K.: An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. *Nature*, 461, 230-235 (2009)

129. Bartsch, H., Voigtsberger, S., Baumann, G., Morano, I. and Luther, H.P.: Detection of a novel sense-antisense RNA-hybrid structure by RACE experiments on endogenous troponin I antisense RNA. *RNA*, 10, 1215-1224 (2004)

130. Kapranov, P., Ozsolak, F., Kim, S.W., Foissac, S., Lipson, D., Hart, C., Roels, S., Borel, C., Antonarakis, S.E., Monaghan, A.P., John, B. and Milos, P.M.: New class of gene-termini-associated human RNAs suggests a novel RNA copying mechanism. *Nature*, 466, 642-646 (2010)

131. Faghihi, M.A. and Wahlestedt, C.: Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol*, 10, 637-643 (2009)

132. Matsui, K., Nishizawa, M., Ozaki, T., Kimura, T., Hashimoto, I., Yamada, M., Kaibori, M., Kamiyama, Y., Ito, S. and Okumura, T.: Natural antisense transcript stabilizes inducible nitric oxide synthase messenger RNA in rat hepatocytes. *Hepatology*, 47, 686-697 (2008)

Key Words: Antisense transcript, Non-coding RNA, Transcriptomics, Genomics, Review

Send correspondence to: Hidenori Kiyosawa, Transdisciplinary Research Integration Center, Research Organization of Information and Systems, National Institute of Genetics, Mishima, Shizuoka, 411–8540, Japan, Tel: 81-55-981-5869, Fax: 055-981-6817, E-mail: hkiyosaw@lab.nig.ac.jp

http://www.bioscience.org/current/vol17.htm