Downloaded from www.genesdev.org on September 4, 2007

Genes & Development

Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum

Kannanganattu V. Prasanth and David L. Spector

Genes & Dev. 2007 21: 11-42
Access the most recent version at doi:10.1101/gad.1484207

References This article cites 370 articles, 157 of which can be accessed free at:
http://www.genesdev.org/cgi/content/full/21/1/11#References

Article cited in:
http://www.genesdev.org/cgi/content/full/21/1/11#otherarticles

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Topic collections Articles on similar topics can be found in the following collections

Chromatin and Gene Expression (80 articles)
Cancer and Disease Models (59 articles)

Notes

To subscribe to *Genes and Development* go to:
http://www.genesdev.org/subscriptions/

© 2007 Cold Spring Harbor Laboratory Press

Downloaded from www.genesdev.org on September 4, 2007

REVIEW

Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum

Kannanganattu V. Prasanth and David L. Spector[1]

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

A large portion of the eukaryotic genome is transcribed as noncoding RNAs (ncRNAs). While once thought of primarily as "junk," recent studies indicate that a large number of these RNAs play central roles in regulating gene expression at multiple levels. The increasing diversity of ncRNAs identified in the eukaryotic genome suggests a critical nexus between the regulatory potential of ncRNAs and the complexity of genome organization. We provide an overview of recent advances in the identification and function of eukaryotic ncRNAs and the roles played by these RNAs in chromatin organization, gene expression, and disease etiology.

Although Jacob and Monod (1961) suggested early on that structural genes encode proteins and regulatory genes produce noncoding RNAs (ncRNAs), the prevailing view has been that proteins not only constitute the primary structural and functional components of cells, but also constitute most of the regulatory control system in both simple and complex organisms. However, recent advances in the fields of RNA biology and genome research have reassessed this "age-old" assumption and provided significant evidence of the importance of RNAs as "riboregulators" outside of their more conventional role as accessory molecules.

Recent large-scale studies of the human and mouse genomes have revealed that although there are ~21,561 protein-coding genes in human and 21,839 in mouse, significantly larger portions of both genomes are transcribed (69,185 gene predictions in human and 71,259 in mouse) (http://www.ensembl.org). Based on such analyses, eukaryotic genomes appear to harbor fewer protein-coding genes than initially expected, and gene number does not scale with complexity as steeply as originally anticipated (Mattick 2004a; Mattick and Makunin 2006). For example, the *Drosophila melanogaster* genome contains only twice as many genes as some bacterial species, although the former is far more complex in its genome organization than the latter. Similarly, the number of protein-coding genes in human and in the nematode *Caenorhabditis elegans* is extremely close (http://www.ensembl.org). Such analyses suggest that protein-coding genes alone are not sufficient to account for the complexity of higher eukaryotic organisms. Interestingly, from genomic analysis it is evident that as an organism's complexity increases, the protein-coding contribution of its genome decreases (Mattick 2004a,b; Szymanski and Barciszewski 2006). A portion of this paradox can be resolved through alternative pre-mRNA splicing, whereby diverse mRNA species, encoding different protein isoforms, can be derived from a single gene (Lareau et al. 2004). In addition, a range of post-translational modifications contributes to the increased complexity and diversity of protein species (Yang 2005).

It is estimated that ~98% of the transcriptional output of the human genome represents RNA that does not encode protein (Mattick 2005). This suggests that these genomes are either replete with largely useless transcription or that these ncRNAs are fulfilling a wide range of unexpected functions in eukaryotic biology (Huttenhofer et al. 2005; Mattick 2005). Recent observations strongly suggest that ncRNAs contribute to the complex networks needed to regulate cell function and could be the ultimate answer to the genome paradox (Mattick 2001, 2003, 2004a,c; Mattick and Gagen 2001). Initially the term ncRNA was used primarily to describe eukaryotic RNAs that are transcribed by RNA polymerase II (RNA pol II) and have a 7-methylguanosine cap structure at their 5′ end and a poly(A) tail at their 3′ end, but lack a single long ORF. However, more recently this classification has been extended to all RNA transcripts that do not have a protein-coding capacity. NcRNAs include introns and independently transcribed RNAs, with the latter accounting for 50%–75% of all transcription in higher eukaryotes (Mattick and Gagen 2001; Shabalina and Spiridonov 2004). Introns account for at least 30% of the human genome, but they have been largely overlooked due to the general assumption that they are rapidly degraded upon pre-mRNA splicing (Mattick 1994, 2005). In mammalian genomes, introns comprise ~95% of the sequence within protein-coding genes. Introns have been suggested to play important roles in nucleosome formation and chromatin organization, alternative pre-mRNA splicing, and as scaffold/matrix-attachment regions (Shabalina and Spiridonov 2004). Intronic sequences have also been shown to harbor independent

[*Keywords*: RNA; disease; intergenic transcripts; noncoding RNA; nuclear regulatory RNAs]
[1]**Corresponding author.**
E-MAIL spector@cshl.edu; FAX (516) 367-8876.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.1484207.

Downloaded from www.genesdev.org on September 4, 2007

transcription units, such as microRNAs, small nucleolar RNAs (snoRNAs), and repetitive elements (Mattick and Makunin 2005).

It is not clear how many ncRNA genes are present in the mammalian genome. The existing catalog of mammalian genes is strongly biased toward protein-coding genes. Novel ncRNA genes are difficult to identify based on sequence analysis due to their sequence divergence across phyla (Pang et al. 2006). The nature of ncRNA genes, including their variation in length (20 nucleotides [nt] to >100 kb), lack of ORFs, and relative immunity to point mutations makes them difficult targets for genetic screens. Analysis of mouse full-length cDNAs revealed that ncRNAs constitute more than one-third of all identified transcripts (Okazaki et al. 2002; Numata et al. 2003; Carninci et al. 2005). Whole human chromosome analysis using oligonucleotide tiling arrays has demonstrated a significantly large number of genes encoding ncRNAs on most of the analyzed chromosomes (Kapranov et al. 2002; Cawley et al. 2004; Kampa et al. 2004; Cheng et al. 2005), many of which show extraordinarily complex patterns of interlaced and overlapping transcription (Carninci et al. 2005; Kapranov et al. 2005). Current estimates of the number of independent transcription units (~70,000) and protein-coding genes (~21,500) in the mammalian transcriptome suggest that ncRNA genes are highly abundant in the genome (Mattick 2004b,c, 2005; Mattick and Makunin 2006; Willingham and Gingeras 2006).

Based on functional relevance, ncRNAs can be subdivided into two classes: (1) housekeeping ncRNAs and (2) regulatory ncRNAs. Housekeeping ncRNAs are generally constitutively expressed and are required for the normal function and viability of the cell. Some examples include transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), small nuclear (snRNAs), snoRNAs, RNase P RNAs, telomerase RNA, etc. These RNAs have been the focus of many reviews (Eddy 2001; Gesteland et al. 2006) and will not be considered further here. In contrast, regulatory ncRNAs or riboregulators include those ncRNAs that are expressed at certain stages of development, during cell differentiation, or as a response to external stimuli, which can affect the expression of other genes at the level of transcription or translation. Several recent excellent reviews have focused on small regulatory RNAs, including small interfering RNAs (siRNAs) and microRNAs (Hannon 2002; He and Hannon 2004; Mattick and Makunin 2005; Zamore and Haley 2005; Petersen et al. 2006), and therefore will not be extensively discussed here, except for their involvement in various diseases. In the present review, we discuss our current understanding of the roles of other noncoding regulatory RNAs in eukaryotic cells and their involvement in gene organization, regulation, and disease etiology.

Roles of RNA in dosage compensation and sex determination: everything needs to be equal

In most animals, the males and females differ in the number of X chromosomes. The expression levels of X-chromosome genes must therefore be equalized in the two sexes, a process referred to as dosage compensation. This can be achieved either by X-chromosome inactivation (XCI) in XX cells or by hyperactivation of the single X chromosome in XY cells. Both of these mechanisms are used by different species and both depend on the expression of regulatory ncRNAs that are key elements of the pathways leading to chromatin remodeling (Lucchesi et al. 2005).

XCI

In mammals, dosage compensation of X-linked gene products between the sexes is achieved by transcriptional silencing of a single X chromosome during early female embryogenesis (Lyon 1961; Plath et al. 2002; Heard and Disteche 2006; Spencer and Lee 2006). Initiation of XCI requires the counting of X chromosomes. XCI follows the "$n - 1$" rule that leads to transcriptional silencing of all but one X chromosome. In female soma, XCI occurs in early development shortly after uterine implantation of the embryo. This form of XCI is called "random" because silencing can take place on either X chromosome (Spencer and Lee 2006). However, in the extraembryonic tissues of some placental mammals, such as rodents, XCI takes place in an "imprinted" manner such that the paternal X (Xp) is always silenced (Takagi and Sasaki 1975). Earlier classical cytogenetics studies suggested that the paternal X only becomes inactivated at the blastocyst stage, accompanying cellular differentiation in the trophoectoderm and primitive endoderm (Takagi et al. 1982). However, recent studies have revealed that the paternal X has already begun to inactivate by the eight-cell stage (Huynh and Lee 2003; Mak et al. 2004; Okamoto et al. 2004; Okamoto and Heard 2006) and this inactivation of Xp initiates following *Xist* RNA coating at the four-cell stage (Okamoto et al. 2004, 2005). Imprinted XCI is also observed in marsupials and is believed to be the earliest form of XCI (Graves 1996). This inactive state is stably maintained through subsequent cell divisions. The *X inactivation center* (*Xic*) is a critical region of 80–450 kb on the X chromosome that controls XCI initiation and spreading (Heard and Disteche 2006; Spencer and Lee 2006). Only the chromosomes carrying the *Xic* sequence are able to induce XCI, even though the "random" and "imprinted" forms of XCI may differ with respect to the requirement of the *Xic* sequences (Okamoto et al. 2005). Interestingly, when *Xic* sequences are inserted into an autosome, the autosome becomes subject to counting, choice, and inactivation (Spencer and Lee 2006).

Of the several long ncRNA genes present in *Xic*, *Xist* (*X-inactive-specific transcript*) has been the most extensively studied ncRNA gene. The *Xist* gene encodes a ncRNA that is associated exclusively with the inactive X chromosome (Fig. 1; Brockdorff et al. 1992; Brown et al. 1992). Although potential ORFs exist in *Xist* RNA, they are short and not conserved between species (Brockdorff et al. 1992; Brown et al. 1992). The gene is conserved between species at the level of its genomic organization

Downloaded from www.genesdev.org on September 4, 2007

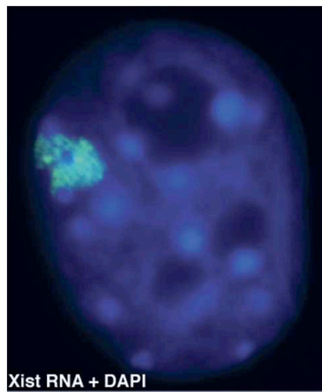

Figure 1. Interphase mouse nucleus showing the localization of *Xist* RNA (green) in the inactive X chromosome. DNA is counterstained with DAPI (blue). (Image provided by Edith Heard, Curie Institute, Paris, France.)

but shows only weak sequence homology, possibly implying a role for its secondary structure. *Xist* ncRNAs are 15–17 kb long in mice, ~19 kb in human, are spliced, polyadenylated, and are restricted to the nuclear compartment. In the female embryo, *Xist* up-regulation on the putative inactive X chromosome (Xi) and RNA coating of this chromosome constitute the first detectable signs of XCI (Morey and Avner 2004). Using inducible *Xist* cDNA transgenes, it was shown that *Xist*-RNA-induced X-chromosome silencing occurs only during early embryonic stem (ES) cell differentiation (Wutz and Jaenisch 2000). However, during initial phases of ES cell differentiation, XCI can be reversed by switching off the *Xist* gene, but subsequently the repressed state becomes locked in and is no longer dependent on *Xist.* This irreversibility of silencing of Xi can be attributed to changes in chromatin modifications observed on the Xi followed by *Xist* RNA coating (Heard and Disteche 2006). The earliest chromatin modifications observed are the loss of histone modifications associated with active chromatin, such as H3K9 acetylation and H3K4 methylation. Subsequently, the X chromosome becomes H4 hypoacetylated and enriched in H3 Lys 27 (H3K27) trimethylation (Plath et al. 2003; J. Silva et al. 2003). H3K27 hypermethylation is accompanied by other chromatin changes, including H3K9 hypermethylation and H4K20 monomethylation as well as H2A K119 monoubiquitylation, and all of these modifications appear concomitantly with the transcriptional silencing of the X-linked genes (Morey and Avner 2004; Heard and Disteche 2006). The inactive X chromosome is also enriched in the histone variant macroH2A, and *Xist* RNA is necessary for its localization to the inactive X (Costanzi and Pehrson 1998). These successive layers of modifications lead to the establishment of silent chromatin and, in turn, lock the inactive X into a stable heterochromatic state throughout the cell cycle. Deletion and transgene analyses have shown that *Xist* is essential for both imprinted and random XCI and affects only the chromosome that transcribes *Xist* RNA (Penny et al. 1996; Marahrens et al. 1997; Wutz and Jaenisch 2000). However, *Xist* alone cannot account for the multiple functions attributed to the *Xic*, such as "counting," as deletion of one *Xist* allele still allows the cell to register the presence of less than one *Xic*, which triggers XCI via the wild-type *Xist* allele (Penny et al. 1996). Interestingly, multiple DNA elements 3′ to *Xist* appear to be involved in counting and choice functions (Heard and Disteche 2006).

Although *Xist* is associated with X-chromosome silencing, its mechanism of action remains unclear. With its noncoding properties, *Xist* could conceivably function through its RNA, either by modulating transcription at the locus, or through organizing chromatin. Several lines of evidence strongly favor the view that *Xist* functions as an RNA (Spencer and Lee 2006). These include (1) the physical association of *Xist* RNA with the inactive X chromosome and the nuclear matrix around the X chromosome. In support of this, recent studies have shown that *Xist* RNA defines a silent nuclear compartment around the future Xi early in the XCI process (Chaumeil et al. 2006; Clemson et al. 2006; Heard and Disteche 2006). (2) Mutations that decrease *Xist* RNA localization generally correlate with reduced silencing (Newall et al. 2001). (3) Expression of *Xist* RNA from an autosome during ES cell differentiation initiates inactivation of the autosome carrying the *Xist* transgene (Lee and Jaenisch 1997). (4) The repeat-A region that contains A-repeats located within intron 1 of *Xist* RNA, which is required for silencing, functions only when placed in the forward (native) orientation (Wutz et al. 2002). The current model suggests that *Xist* RNA initiates silencing by binding to specific silencing factors, recruiting those silencing proteins to the *Xic*, and subsequently propagating those factors along the X chromosome as the RNA itself spreads throughout the chromosome (Spencer and Lee 2006).

Other than *Xist* RNA, the *Xic* region in mouse also harbors many other ncRNA genes including *Tsix, Xite*, and *Jpx/Enox*, several of which are integral to the regulation of XCI. *Tsix* negatively regulates the expression of *Xist* RNA and is transcribed in an antisense orientation relative to *Xist.* Like *Xist*, *Tsix* lacks a conserved ORF and is found only in the nucleus (Lee et al. 1999). In undifferentiated female ES cells, *Xist* and *Tsix* are coexpressed on both X chromosomes, although the *Tsix* levels are in 10- to 100-fold molar excess over *Xist* RNA (Shibata and Lee 2003). However, a recent study suggests that *Xist* is expressed at an extremely low level prior to XCI and that *Tsix* is the major RNA component detected at the *Xist*/*Tsix* locus in undifferentiated ES cells (Sun et al. 2006). At the onset of cell differentiation, *Tsix* becomes asymmetrically expressed: Whereas *Tsix* expression persists transiently on the future active X (Xa), expression ceases on the future inactive X (Xi). The loss of *Tsix* expression on the future Xi enables the up-regulation and spread of *Xist* RNA along the chromosome. The persistence of *Tsix* on the future Xa enables that X to remain active. Once the window for XCI has passed, *Tsix* is also turned off on the Xa. These results suggest that by controlling the fate of *Xist* and therefore the X chromosome, *Tsix* acts as a binary switch for XCI. The reason for

Downloaded from www.genesdev.org on September 4, 2007

this sudden reciprocal expression profile of *Xist* and *Tsix* remains unknown. Interestingly, two recent studies have revealed that the *Xics* transiently colocalize, via the *Tsix* region, during the onset of XCI, at the time when counting and choice are thought to occur (Bacher et al. 2006; Xu et al. 2006). This "cross-talk" between the *Xics* is thought to be required for the exchange of information between *Xist*/*Tsix* that ultimately results in the monoallelic down-regulation of *Tsix* and up-regulation of *Xist* on the inactive X chromosome (Heard and Disteche 2006). Several mechanisms have been proposed to explain how *Tsix* regulates *Xist* (Spencer and Lee 2006). These include (1) a DNA-based mechanism in which DNA sequences at *Tsix* bind transcription factors that then repress the *Xist* promoter at long range, or *Tsix* could also compete with *Xist* for an enhancer or any other regulatory sequence; (2) a transcription-based mechanism, where antisense transcription across the *Xist* promoter could interfere with the ability of the *Xist* promoter to fire by affecting chromatin modification or transcription factor binding; (3) *Tsix* RNA itself could recruit repressive factors or could form duplex RNA with *Xist* that would either facilitate the degradation of *Xist* RNA or prevent binding of necessary silencing factors to *Xist* RNA. Recent studies have provided clues that suggest either *Tsix* transcription or *Tsix* RNA itself has a role in *Xist* RNA regulation (Spencer and Lee 2006). It has been observed that overexpression of *Tsix* always results in an active X in *cis* (Luikenhuis et al. 2001; Stavropoulos et al. 2001). Furthermore, when *Tsix* RNA is prematurely truncated before it crosses into the *Xist* gene, *Tsix* no longer functions as a repressor of *Xist,* and XCI invariably occurs on the mutated X (Shibata and Lee 2004). It was also proposed earlier that the modulation of *Xist* chromatin structure might play a role in how *Tsix* regulates *Xist* (Navarro et al. 2005; Sado et al. 2005). Interestingly, a recent study has suggested that up-regulation of *Xist* RNA observed on the future inactive X is not due to the increased stability of the *Xist* transcript as suggested earlier but is regulated by *Tsix* (Panning et al. 1997; Sheardown et al. 1997; Sun et al. 2006). Lee and colleagues (Sun et al. 2006) reported that *Tsix* down-regulation on the future inactive X induces a transient heterochromatic state to *Xist*, followed by high levels of *Xist* expression. This heterochromatic state adopted by the *Tsix*-deficient chromosome in pre-XCI cells persisted through XCI establishment and reverted to a euchromatic state during XCI maintenance (Sun et al. 2006).

The mouse *Xic* harbors yet another functional ncRNA gene, called *Xite* (*X-inactivation intergenic transcription elements*). *Xite* is transcribed at low levels, on the order of 10- to 60-fold less than *Tsix* levels in mouse ES cells. Although there is some bidirectional transcription, the majority of the transcripts are oriented in the same direction as *Tsix.* Deleting *Xite* results in preferential silencing of the X in *cis*, thereby skewing the normally random probability that any one X would be chosen as the silent one (Ogawa and Lee 2003). *Xite* action does not appear to depend on the RNA per se, because truncation of the RNA does not produce any obvious phenotype, suggesting that transcription from the region could be more important. The monoallelic expression of *Xist*, at least in mice, is controlled by complex regulation of *Tsix* and *Xite* as well as *cis*-regulatory sequences located in the 3′ region of *Xist* (Heard and Disteche 2006). In the current model, *Xite* works together with *Tsix* to designate the Xa where transcription from *Xite* acts as an enhancer for *Tsix* by promoting the persistence of *Tsix* expression during cell differentiation; this in turn prevents the up-regulation and spread of *Xist* RNA along the chosen Xa (Spencer and Lee 2006). *Ftx* is another ncRNA gene located ~150 kb upstream of mouse *Xist.* In mouse and humans, the 5′ regions of *Ftx* are well conserved and contain CpG islands at positions corresponding to the cDNA start sites and are transcribed in opposite orientation relative to *Xist*/*XIST* genes (Chureau et al. 2002). Future investigation of a less-characterized ncRNA gene *Jpx* (*Enox*) found around the *Xic* may also show that this gene participates in the regulatory events of XCI (Spencer and Lee 2006).

X-chromosome hyperactivation in Drosophila

Unlike the situation in mammals, dosage compensation in *Drosophila* is achieved by a twofold up-regulation of transcription of genes on the single X chromosome present in males (Kelley 2004). Intriguingly, the fly dosage compensation system also involves multiple ncRNAs: *roX1* and *roX2* (*RNA on the X*). These RNAs are members of the dosage compensation complex (DCC), a huge RNA–protein complex that binds to hundreds of sites along the male X chromosome in a highly reproducible, banded pattern (Meller 2000; Meller et al. 2000). In addition to *roX1* and *roX2* RNAs, the DCC also contains a specific set of proteins that include MLE (*maleless*); MSL1, MSL2, and MSL3 (*male-specific lethal 1, 2*, and *3*, respectively); and MOF (*males absent on the first*). Mutations in these genes result in male-specific lethality of larvae, and their products are collectively termed MSL proteins. A characteristic feature of the up-regulated X chromosome is the specific acetylation of histone H4 at Lys 16 (H4Ac16) (Akhtar 2003).

The two *roX* genes are transcribed from the X chromosome, produce polyadenylated nuclear retained transcripts, and are expressed only in male adult flies (Fig. 2). The *roX* RNAs are functionally redundant even though they have very little sequence homology and are distinct in size (3.7 kb for *roX1* RNA vs. 0.5–1.2 kb for *roX2* RNA) (Meller and Rattner 2002). Deletion of either *roX* gene has no effect on males. However, deletion of both results in male lethality. The MSL-binding pattern on the X chromosome is drastically disrupted in the *roX1 roX2* double-mutant males, suggesting that *roX* RNAs are important for correctly targeting the MSL complex to the X (Meller and Rattner 2002). The *roX* genes could be performing two distinct and separable functions in dosage compensation. First, *roX* RNAs constitute indispensable elements of the DCC responsible for chromatin modifications. Second, the genes themselves provide strong chromatin entry sites for the MSL complex, possibly to

Downloaded from www.genesdev.org on September 4, 2007

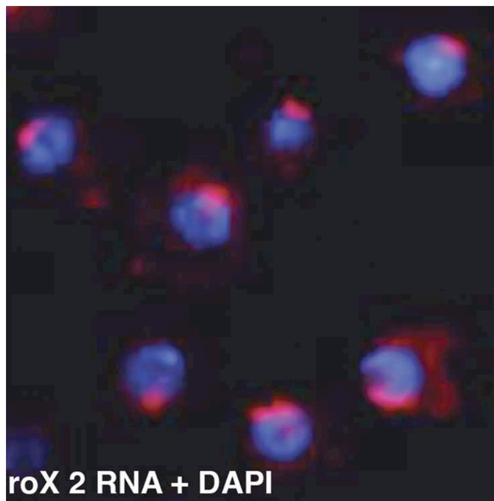

Figure 2. *RoX2* ncRNA (red), visualized by RNA fluorescence in situ hybridization, is localized to the active X chromosome in *Drosophila* male SL2 tissue culture cells. DNA (blue) is counterstained with DAPI. (Image provided by Polina Gordadze and Mitzi I. Kuroda, Brigham and Women's Hospital, Boston, MA, USA.)

ensure rapid recruitment of the MSL proteins for *roX* RNA binding. The current model suggests that there are different DNA recognition elements on the X chromosome that have different affinities for the MSL complex; high, intermediate, or weak. High-affinity *cis* elements, such as within the *roX* genes, would not require additional *cis*-elements for recruiting MSL complexes, and this interaction is strengthened by *roX* RNA. Intermediate and weak-affinity *cis*-elements might require several *cis*-elements for robust binding, resulting in the ability to attract partial MSL complexes (Oh et al. 2004).

The targeting mechanisms of DCC to X chromosomes between mammals and *Drosophila* show some superficial similarities. In both cases, ncRNAs are required for targeting the correct chromosome for regulation. Furthermore, in each case there is evidence for spreading of the DCC process long distances along the chromosome from the sites of synthesis of those ncRNAs. However, the major differences between the mammalian and *Drosophila* systems is that in mammals, the DCC is involved in the inactivation of one of the X chromosomes, whereas in *Drosophila*, it results in the hyperactivation of the single X. *Drosophila* and mammals also differ in that *Xist* is limited to its action in *cis,* while *roX* RNAs and the MSL complex can also clearly act in *trans* (Oh et al. 2004; Heard and Disteche 2006).

Although there is significant evidence to show that ncRNAs are the major effectors of dosage compensation, the molecular basis of how they regulate these processes is still not clearly understood, and the future is likely to reveal many exciting solutions.

Male hypermethylated (MHM) region in birds

In birds, sex determination and differentiation depend on the sex chromosomes Z and W. Males have two Z chromosomes, whereas females are determined by the ZW karyotype. One of the genes proposed to play a role in sex determination in birds is a homolog of human *DMRT1* (doublesex and mab-3-related transcription factor) implicated in testis differentiation. *DMRT1* has been mapped to the Z chromosome, and its elevated expression in males has been found to correlate with testis development (Smith et al. 2003). A MHM region was identified in the Z chromosome in the vicinity of the *DMRT1* gene, and the CpG islands in this region are hypermethylated only in males. However, in females, the MHM region is hypomethylated, and transcription from this region produces ncRNAs (the longest transcripts are ~9.5 kb), most of which are nonpolyadenylated and accumulate at or very close to the sites of transcription and close to the *DMRT1* locus. The female-specific MHM ncRNAs are suggested to play a role as transcriptional repressors of the *DMRT1* locus similar to the role played by *Xist* RNA in XCI (Teranishi et al. 2001; Szymanski and Barciszewski 2003).

Roles of ncRNAs in genomic imprinting: one is enough

Diploid organisms usually express both alleles of an active gene. However, in marsupial and placental mammals, some genes express only one of the alleles, a phenomenon termed "genomic imprinting." Genomic imprinting is a process whereby the expression of an allele depends on whether it is derived from the mother or father (Bartolomei and Tilghman 1997; Verona et al. 2003). Genomic imprinting was first discovered on the X chromosome, where Sharman and colleagues (Richardson et al. 1971; Sharman 1971) described a form of XCI in marsupials in which the paternal X chromosome is preferentially silenced. Other than in mammals, genomic imprinting has also been identified in angiosperm plants and in a few insects (Braidotti et al. 2004). Recent studies have shown that >100 genes are imprinted in mammals, either "paternally imprinted" (the gene is silent on the paternal allele) or "maternally imprinted" (the gene is silent on the maternal allele). The imprinted genes generally exist in clusters on various chromosomes, suggesting that the mechanism to control imprinted expression acts on the chromosomal domains rather than on individual genes. Interestingly, these imprinted clusters often are associated with imprinted ncRNA genes. Expression of the ncRNA from one of the alleles often correlates with the repression of the linked protein-coding gene on the same allele (O'Neill 2005). This reciprocal parental-specific expression of imprinted mRNAs and ncRNAs has long been suggested to indicate that ncRNAs play a role in silencing the mRNA genes in an imprinted cluster (Pauler and Barlow 2006). Some of the imprinted loci in mammalian cells where the presence of ncRNAs is well documented are described below.

IGF2/H19 *locus*

The *Igf2/H19* domain is perhaps the best characterized of any autosomally imprinted locus (human 11p15.5 and

Downloaded from www.genesdev.org on September 4, 2007

mouse distal 7b). The first imprinted ncRNA locus to be discovered, the *H19* gene produces a spliced and polyadenylated ncRNA transcript of ~2.3 kb that is expressed only from the maternal allele (Brannan et al. 1990). *H19* is the reciprocally imprinted partner of *Igf2* (*insulin-like growth factor*), and *Igf2* is expressed only from the paternal allele. Mutations disrupting the imprinted expression of *Igf2* underlie a substantial proportion of cases of congenital growth disorder. Interestingly, in the *Igf2*/*H19* domain, imprinting is achieved through "enhancer competition" mediated by a set of chromatin insulators. *Igf2* and *H19* share a set of enhancers, but only one gene can engage the enhancer at any time and is regulated by an insulator sequence that lies just upstream of the *H19* promoter (Webber et al. 1998; Kanduri et al. 2000; Kaffer et al. 2001). On the maternal chromosome, the insulator sequence is not methylated and, therefore, binds CCCTC-binding factor (CTCF), a vertebrate insulator protein. Binding of CTCF prevents the enhancers from engaging the *Igf2* gene and together with the enhancers also *trans*-activates *H19.* However, on the paternal chromosome, the insulator sequence is methylated and, therefore, cannot bind the methylation-sensitive CTCF, allowing the enhancer to engage *Igf2* (Engel and Bartolomei 2003). In this way, the insulator sequence upstream of *H19* comprises an "imprinting center" that regulates the reciprocal expression of *H19* and *Igf2*. Although *H19* is conserved among mammals and highly expressed in embryos, studies carried out over the last 15 yr indicate that the *H19* transcript itself has no apparent role in the imprinted expression of its neighboring genes (Jones et al. 1998) and is also not necessary for normal development in mice (Ripoche et al. 1997). The chromosomal region containing *H19* has also been associated with tumor suppressor activity, and the expression pattern of *H19* RNA in several cancer cell types differs from neighboring nonmalignant cells (see "Regulatory RNAs Implicated in Complex Diseases: Dark Side of RNA" below). In addition to *H19*, other ncRNAs emanating from the *Igf2*/*H19* region have been identified, some of which show imprinting while others are expressed biallelically; however, their functional significance has yet to be determined (Moore et al. 1997; Drewell et al. 2002b).

KCNQ1 *locus*

As with the closely linked *Igf2*/*H19* cluster, the *KCNQ1* locus is closely associated with human Beckwith-Wiedemann syndrome (BWS), a syndrome characterized by parental asymmetric overgrowth, enlarged tongue, and cancer such as Wilms' tumor (Szymanski and Barciszewski 2003; O'Neill 2005). The inheritance of BWS is exceptionally complex because the etiology of the disease involves multiple genes in both the *KCNQ1* and the *Igf2*/*H19* domains. Interestingly, almost all of the imprinted genes in the *KCNQ1* domain are maternally expressed except the paternally expressed ncRNA gene *Kcnq1ot1* (*Lit1*), the antisense counterpart of *Kcnq1* (Mitsuya et al. 1999; Umlauf et al. 2004). The antisense *Kcnq1ot1* gene appears to be critical for establishing the imprinted profile of the nearby genes (Mancini-Dinardo et al. 2006). Recent studies suggest that the *Kcnq1ot1* RNA does so by the recruitment of chromatin changes to the imprinted domain, including H3K9 methylation and H3K27 methylation (Lewis et al. 2004; Umlauf et al. 2004). The *Kcnq1ot1* promoter lies within a differentially methylated region of the *Kcnq1* gene body and is now known to make up the imprinting center for the BWS domain (Spencer and Lee 2006). Deleting the *Kcnq1ot1* CpG island (5′ end) results in loss of imprinting in mice, and either the *Kcnq1ot1* RNA or transcription through its entire length is required in *cis* for imprinting of neighboring genes (Cleary et al. 2001; Thakur et al. 2004; Mancini-Dinardo et al. 2006). A transgenic mouse producing a truncated *Kcnq1ot1* transcript exhibited correct imprinting but does not result in silencing any of the flanking mRNA genes in the imprinted cluster (Mancini-Dinardo et al. 2006; Pauler and Barlow 2006). Interestingly, the most common abnormalities in BWS are epigenetic, involving abnormal methylation of *H19* or *Kcnq1ot1.* Recently, microdeletions either in the *H19* or *Kcnq1ot1* gene have been shown to be associated with BWS, providing genetic confirmation of the importance of this chromosomal region for the disease (Costa 2005).

Igf2r *(*insulin-like growth factor type-2 receptor*)*/Air *(*Antisense Igf2r RNA*)*

The *Igf2r*/*Air* locus (proximal chromosome 17) in mice provides yet another example of ncRNA regulation within imprinted loci. A differentially methylated region-2 (DMR2) within the second intron of *Igf2r* constitutes a critical, bidirectional element controlling silencing of the paternal allele of three protein-coding imprinted genes, *Igf2r*, *Slc22a2*, and *Slc22a3* (Zwart et al. 2001). DMR2 resides in a promoter that drives the transcription of a nonprotein-coding antisense transcript, *Air*, which partially overlaps with *Igf2r. Air* is an ~108-kb, capped, polyadenylated, ncRNA and is transcribed exclusively by RNA pol II from the paternal allele (Wutz et al. 1997; Braidotti et al. 2004; Seidl et al. 2006). The majority of *Air* transcripts evade cotranscriptional splicing resulting in mature unspliced, highly unstable nuclear transcripts (Seidl et al. 2006). Like *Kcnqlot1*, the *Air* gene is responsible for the bidirectional silencing of neighboring genes in *cis,* as deleting the *Air* CpG island results in loss of parental silencing across the entire domain (Wutz et al. 1997; Zwart et al. 2001). The silencing of these three genes depends on the unmethylated CpG islands and transcription of *Air* RNA. Because *Air* RNA does not overlap with two of the three imprinted genes in the domain (*Slc22a2* and *Slc22a3*), *Air* RNA cannot work through double-stranded RNA (dsRNA) mechanisms, but because truncating *Air* RNA leads to a disruption of imprinting, its transcription and/or the RNA itself may be required for imprinting (Sleutels et al. 2002; Spencer and Lee 2006). A suggested mechanism of *Air* action involves two steps. First, *Air* expression results in the silencing of the overlapping *Igf2r* by promoter occlusion or *cis*-acting RNA interference (RNAi). This could

Downloaded from www.genesdev.org on September 4, 2007

result in an induction of the silent chromatin state that would spread and shut off flanking genes. However, studies by Barlow and colleagues (Sleutels et al. 2003) showed proper imprinting of *Slc22a2*, *Slc22a3*, and also *Air* in mice that lack *Igf2r*, suggesting that the antisense mechanism followed by spreading of silencing may not be the only mechanism responsible for *Igf2r/Air* locus imprinting. Alternatively, *Air RNA* could recruit chromatin modifier proteins to specific regions of the imprinted locus in a manner similar to the role suggested for *Xist* RNA (Sleutels et al. 2003). Consistent with this, *Igf2r* exhibits allele-specific histone modifications (Fournier et al. 2002). However, RNA FISH analysis using specific probes against *Air* RNA did not show coating by *Air* RNA of the imprinted chromosomal region (Braidotti et al. 2004).

Prader-Willi/Angelman syndrome (PWS/AS) locus

PWS AS are the result of disrupted expression of imprinted genes covering a >4-Mb region of human 15q11–13 (mouse proximal 7). The PWS/AS locus in human provided the first example of an imprinted disorder when it was discovered that uniparental disomies (the inheritance of both chromosome copies from the single parent) of chromosome 15 results in an assemblage of congenital problems (O'Neill 2005). Maternal disomies result in PWS, whereas the paternal disomies result in AS. PWS is exemplified in newborns by hypotonia, hypogonadism, and various mental retardation and feeding difficulties, followed later in childhood by hyperphagia (Cassidy et al. 2000). PWS is a continuous gene disorder manifested by loss of expression of a group of paternally transcribed protein-coding genes including *SNURF/SNRPN*, *MKRN3*, *MAGEL2*, and *ZNF127* (O'Neill 2005). *IPW* (Imprinted in Prader-Willi) was isolated as a novel imprinted ncRNA gene from the PWCR (Prader-Willi chromosome region) that produces a spliced and polyadenylated ncRNA (Wevrick et al. 1994). The same locus also codes for another ncRNA gene, *ZNF127AS*, an antisense gene to *ZNF127* expressed in brain and lungs (Jong et al. 1999). AS is characterized by ataxic gate, jerky arm movements, inappropriate laughter, and severe mental retardation (Williams et al. 1995). Loss-of-function mutations in a maternally transcribed gene at this locus, *UBE3A*, can cause AS (Albrecht et al. 1997; Kishino et al. 1997). The paternal silencing of *UBE3A* is confined to specific brain subregions; elsewhere it is biallelically expressed (Rougeulle et al. 1997; Vu and Hoffman 1997). Additionally, there is paternal-specific expression of a large, alternatively spliced antisense transcript (*UBE3A-ATS*), spanning ~450 kb in human and ~1 Mb in mice. Deleting the 5′ end of this long antisense transcript results in reduced expression of *UBE3A* on the paternal chromosome (Chamberlain and Brannan 2001). Although no role has been ascribed to the large *UBE3A* antisense transcripts, it has been proposed that these RNAs may be directly linked to the etiology of the diseases (Rougeulle et al. 1998; Runte et al. 2004). A second maternal-specific transcript from this region, *ATP10C*, has also been implicated in the AS phenotype (Meguro et al. 2001). The PWS/AS locus also contains several clusters of snoRNAs (*C/D-box snoRNAs*) expressed exclusively from the paternal chromosome. Interestingly, many of these snoRNA genes that overlap *UBE3A* on the opposite strand were shown to be overexpressed in AS patients (Runte et al. 2001).

GNAS *locus*

Transcription of genes at the *GNAS* imprinted locus (human 20q13 and mouse distal 2) is exceptionally complex. The core gene of this locus is *GNAS*, which is expressed ubiquitously and biallelically in all but a few tissues. It encodes Gsα, the α-subunit of the heterotrimeric G-protein complex. Constitutive activating mutations in Gsα give rise to McCune-Albright syndrome, characterized variably by café-au-lait spots, gonadotropin-independent sexual precocity, and fibrous dysplasia of bone (Schwindinger et al. 1992). In certain hormone targeted tissues (renal proximal tissues, gonads, and thyroid in humans), *GNAS* is transcribed predominantly from the maternal allele. *NESP55*, encoding a chromogranin-like neurosecretory protein, is also maternally expressed. Unusually, *NESP55* incorporates exons 2–13 of *GNAS* into its 3′ untranslated region (UTR). The ncRNAs transcribed from this locus includes *NESPAS*, a spliced antisense transcript, and a truncated ncRNA transcript expressed from the *GNAS* locus by alternative promoter usage. A recent report implicates a possible role for the *NESPAS* transcript in the transcriptional control of *GNAS* (Bastepe et al. 2005). *NESPAS* RNA expression could repress *NESP55* by promoter occlusion, localized heterochromatinization, or competition for shared transcription factors (Wroe et al. 2000).

Study of the molecular elements that combine to initiate and maintain the imprint and translate it into monoallelic expression has suggested a critical role of ncRNAs in governing gene silencing. Better insight into the mechanism of ncRNA action on the imprinted loci will provide an important paradigm for understanding genomic imprinting.

Intergenic transcripts: sense in reading between the genes

A large proportion of transcripts from eukaryotic genomes correspond to intergenic transcripts and antisense transcripts. The intergenic transcription units produce ncRNAs of variable sizes that are not well conserved across the phyla (Babak et al. 2005). Although the exact functions of these RNAs have not been validated, their functions are likely linked to transcription-dependent mechanisms rather than being RNA-dependent per se. There are already several examples of intergenic transcription associated with developmentally regulated genes, which play important roles in the coordination of gene expression. Several of the more well-documented intergenic transcription sites include the following.

Downloaded from www.genesdev.org on September 4, 2007

Mammalian β-globin *locus*

In humans, the 70-kb *β-globin* locus consists of five erythroid-specific genes; embryonic (ε), fetal ($^{G}\gamma$ and $^{A}\gamma$), and adult (δ and β), whose expression is under the control of the β-LCR (locus control region). Analysis of nascent transcripts from the *β-globin* gene cluster revealed that both intergenic regions and LCR constitutively produce specific ncRNAs (Ashe et al. 1997). Both LCR and intergenic transcripts originate from the same strand as other globin genes and are retained in the nucleus (Ashe et al. 1997). Expression of ncRNA transcripts from the LCR and intergenic regions are restricted primarily to erythroid cells. Interestingly, transient expression of globin genes in nonerythroid cells can induce transcription from the intergenic region without activating the protein-coding domains (Ashe et al. 1997). An explanation for the production of intergenic trancripts from the LCR has been suggested by a "tracking model." According to this model, erythroid-specific and ubiquitous transcription factors and cofactors form complexes with the LCR and track along the locus. When this transcription complex encounters the basal transcription machinery, located at the promoter, transcription of the gene is initiated (Q. Li et al. 2002). During this process, there is a high probability that intergenic transcripts would arise from the cryptic start sites along the locus. It has been proposed that these intergenic transcripts might facilitate the recruitment of *trans*-acting factors and RNA pol II to the promoters of *globin* genes via this tracking mechanism (Tuan et al. 1992). Alternatively, intergenic transcription may be required for the establishment and maintenance of an open chromatin conformation within the *globin* locus (Gribnau et al. 2000; Plant et al. 2001). However, the persistence of DNase I hypersensitivity following deletion of the LCRs in cell lines argues against this role (Epner et al. 1998; Reik et al. 1998). Similarly, studies by Haussecker and Proudfoot (2005) did not observe a positive correlation between intergenic transcript abundance and chromatin activation and/or *globin* gene expression. Instead, this study suggested that intergenic transcription at the *β-globin* locus mediates the formation of silent chromatin in the absence of erythrocyte-specific transcription factors (Haussecker and Proudfoot 2005).

IL-4/IL-13 *gene cluster*

During differentiation of naive CD4^{+} precursors to T helper 1 (Th1) or Th2 effector cells, several epigenetic changes occur in a lineage-specific manner at the *IFN*γ or *IL4*/*IL13* loci. Upon activation, a subset of Th2 cells involved in cell-mediated immune responses express *IL-4* and *IL-13* genes located in tandem on human chromosome 5q (chromosome 11 in mouse) (Frazer et al. 1997). This cluster is flanked by two constitutively expressed genes: *Rad50* and *Kif3a.* Transcription analysis from this intergenic region in CD4^{+} T cells has revealed the presence of a 130- to 260-nt polyadenylated nuclear retained ncRNA. Studies in a mouse transgenic model have revealed that the intergenic transcription is restricted to tissues and lineages in which *IL-4* and *IL-13* are expressed and is up-regulated upon Th2 differentiation (Rogan et al. 2004). However, these intergenic transcripts are constitutively expressed even in the absence of active *IL* genes, implying that they are derived from independent transcription units. Although the role of these intergenic transcripts is not clear, one possible explanation is that they result from the chromatin remodeling activity at this locus (Takemoto et al. 2000). Consistent with this idea, the differentiation of Th2 cells was found to be associated with hyperacetylation of histone H3 and hypomethylation of the CpG islands (Yamashita et al. 2002). Another example of integenic transcription in a lineage-specific gene cluster has been described at the MHC class II locus (Masternak et al. 2003).

Intergenic transcripts from the Dlx-5/6 *region*

Vertebrate *Dlx* genes are members of the homeodomain protein family that play critical roles in differentiation and migration of neurons as well as craniofacial and limb patterning during development (Feng et al. 2006 and references therein). The *Dlx* genes are expressed in bi-gene clusters, and conserved intergenic enhancers have been identified for the *Dlx-5/6* and *Dlx-1/2* loci (Zerucha et al. 2000; Ghanem et al. 2003). One of the two conserved intergenic regions from mouse, the *Dlx-5/6* region transcribes two ncRNAs, *Evf-1* and *Evf-2*, the latter being the alternatively spliced form of *Evf-1* (Kohtz and Fishell 2004; Feng et al. 2006). *Evf-1* is a 2.7-kb polyadenylated RNA, and its expression is developmentally regulated (Kohtz and Fishell 2004). The *Evf-2* ncRNA (3.8 kb) specifically cooperates with the homeodomain protein Dlx-2 to increase the transcriptional activity of the *Dlx-5/6* enhancer region in a target- and homeodomain-specific manner. Interestingly, a stable complex containing the *Evf-2* ncRNA/Dlx-2 homeodomain protein forms in vivo in the nucleus (Feng et al. 2006). Together, these data suggest that the *Evf-2*/Dlx-2 complex stabilizes the interaction between Dlx-2 and target *Dlx-5/6* enhancer sequences to increase transcriptional activity. The role of *Evf-2* as a transcriptional activator suggests the possibility that a subset of such vertebrate ultraconserved regions may function at the RNA level as key developmental regulators.

Bithorax complex (BX-C) in Drosophila

In *Drosophila*, the homeotic genes encoded by the BX-C are involved in specifying the segmentation of the embryo and determining the body plan (Lewis 1978). The correct spatial and temporal expression of the three protein-coding genes *Ultrabithorax* (*Ubx*), *Abdominal-A* (*Abd-A*), and *Abdominal-B* (*Abd-B*) is crucial for the development of thoracic and abdominal segments. The expression pattern of *Abd-A* and *Abd-B* depends on an array of regulatory elements located in the intergenic regions between these genes, including seven genetically

Downloaded from www.genesdev.org on September 4, 2007

defined *infra-abdominal* (*iab-2–8*) domains, and mutations in this region are associated with developmental defects affecting abdominal segments (Sanchez-Herrero and Akam 1989). The *iab*s are transcribed exclusively in the embryos. A systemic examination of the distribution of these intergenic transcripts from the *iab* regions revealed that they show highly specific localization along the anterior–posterior axis of the blastoderm embryo and the transcripts are restricted to the nucleus (Bae et al. 2002). The intergenic transcripts originating from *iab-4* revealed 1.7-kb and 2.0-kb polyadenylated ncRNAs that are transcribed in the opposite direction to *Abd-A* (Cumberledge et al. 1990). Alteration of transcription in one *iab* subdomain induces a homeotic transformation of the more posterior segment under its control, suggesting that intergenic transcription plays a crucial role in *iab* activity (Drewell et al. 2002a). Intergenic transcription from the *iab* regions has also been proposed to play a role in the activation of *cis*-regulatory elements by interfering with the Polycomb-repressing complex, responsible for silencing the homeotic genes (Bender and Fitzgerald 2002; Hogga and Karch 2002). The *iab-4* region contains a single ~100-nt pre-miRNA hairpin structure that encodes two stable miRNAs: *mir-iab-4-5p* and *mir-iab-4-3p* (Aravin et al. 2003). Recent studies revealed that these miRNAs regulate *Ubx* activity in vivo (Stark et al. 2003; Grun et al. 2005; Ronshaugen et al. 2005).

Intergenic transcription within the BX-C is not limited to the *iab* regions but also has been reported for the bithoraxoid (*bxd*) region (Lipshitz et al. 1987). This region exhibits active transcription twice: once early in embryogenesis and once in later larval and adult stages. The early transcripts (1.1–1.3 kb, are processed from a 26-kb precursor) appear to be ncRNAs, whereas the late transcripts (0.8 kb) can be translated to produce a protein (Lipshitz et al. 1987). Recently, an elegant study by Sauer and colleagues (Sanchez-Elsner et al. 2006) provided direct evidence of the role of intergenic transcripts from the *Ubx* region in epigenetic activation of gene expression. The *Ubx* locus contains multiple *cis*-regulatory elements known as trithorax response elements (TRE) that recruit transcriptional activators such as the trithorax group (trxG) of epigenetic regulators. Interestingly, the same DNA elements can also act as repressor-binding sites, Polycomb response elements (PRE), and facilitate the recruitment of members of the Polycomb (PcG) complex. It has previously been shown that intergenic transcription of ncRNAs from TRE/PRE elements switches a silent PRE to a TRE, which indicates that TRE/PRE transcription plays an important role in epigenetic activation (Lipshitz et al. 1987; Rank et al. 2002; Schmitt et al. 2005). Recent studies by Sanchez-Elsner (Sanchez-Elsner et al. 2006) further showed that these intergenic transcripts from the TRE at the *Ubx* locus mediate transcriptional activation of *Ubx* by recruiting the epigenetic regulator Ash1 to the TRE elements. Ash1 is a histone methyltransferase (HMT) that promotes transcriptional activation by trimethylating H3K4, H3K9, and H4K20 (Beisel et al. 2002) and is essential for the tissue-specific expression of *Ubx* (Beisel et al. 2002 and references therein). Therefore, intergenic transcripts derived from the TRE locus mediate the recruitment of Ash1 to the TRE DNA elements of *Ubx*. These ncRNA transcripts serve as an intermediary between the TRE DNA elements and Ash1 protein (Sanchez-Elsner et al. 2006). These data further support a model in which an intergenic ncRNA transcribed from the TRE of *Ubx* is retained at the TRE through DNA–RNA interactions and plays an important role in providing an RNA scaffold that is recognized by Ash1.

SRG1 *in* Saccharomyces cerevisiae

Unlike the above examples in which intergenic transcription is involved in the transcriptional activation of the corresponding region, studies in the budding yeast *S. cerevisiae* have revealed the role of intergenic transcription in transcriptional repression (Martens et al. 2004, 2005). Transcription of the intergenic ncRNA gene *SRG1* (*SER3 regulatory gene 1*) across the promoter of the adjacent *SER3,* a serine biosynthetic gene, represses the transcription of *SER3* by transcriptional interference (Martens et al. 2004). *SRG1* transcription is regulated by serine such that in the presence of serine, the serine-dependent activator Cha4 binds to the *SRG1* promoter and activates its transcription, thereby negatively regulating the expression of *SER3* (Martens et al. 2005). These studies demonstrate an example where intergenic transcription provides a mechanism for a single protein, Cha4, to simultaneously activate and repress opposing pathways.

The evergrowing list of intergenic transcripts located mostly in the nonprotein-coding regions of the genome has highlighted the importance of intergenic transcription in regulating gene activity. This further highlights the fact that the high proportion of nonprotein-coding regions in the eukaryotic genome is probably not due to the accumulation of nonsense DNA but rather represents the evolution of more complicated gene regulatory mechanisms (Schmitt and Paro 2004).

Natural antisense transcripts (NATs): new players in the gene regulatory network

Computational analysis of data from large-scale sequencing projects has revealed a surprising abundance of NATs in several eukaryotic genomes (Lehner et al. 2002; Lavorgna et al. 2004). More than 2500 NATs have been identified in human of which >1600 are predicted to be true NATs (Yelin et al. 2003). Recent genome-wide analyses suggest that as much as 15%–25% of human genes might be involved in antisense transcription (http://www.narna.ncl.ac.uk). Similar analyses in other organisms including mouse have revealed a large number of NATs (Kiyosawa et al. 2003, 2005; Lavorgna et al. 2004; Katayama et al. 2005). NATs are RNAs containing sequences that are complementary to other endogenous RNAs. They can be transcribed in *cis* from opposing DNA strands at the same genomic locus (*cis*-NATs) or in

Downloaded from www.genesdev.org on September 4, 2007

trans from separate loci (*trans*-NATs). In human tissues, the sense–antisense pairs tend to be coexpressed and/or inversely expressed more frequently than expected by chance, and this expression pattern tends to be evolutionarily conserved (Chen et al. 2005). NATs have been implicated in many levels of eukaryotic gene regulation including translational regulation, genomic imprinting, RNAi, alternative splicing, XCI, RNA editing, and gene silencing (Kumar and Carmichael 1997; Lavorgna et al. 2004). Even though the eukaryotic genome contains a large number of NATs, our understanding of how antisense transcription regulates gene expression remains largely incomplete. The regulation of gene expression by NATs can occur through multiple mechanisms, as shown below.

Transcriptional interference

Transcription by RNA pol II involves both large protein complexes and the unwinding of the duplex DNA. It is unlikely that two overlapping transcriptional units could be transcribed concomitantly by the RNA pol II machinery. Such effects have been well studied with respect to the *GAL10* and *GAL7* genes in *S. cerevisiae* (Prescott and Proudfoot 2002). When arranged convergently, but not overlapping, both genes are transcribed at normal levels. However, when the two transcription units overlap, steady-state mRNA levels are severely reduced due to an inhibition of transcription elongation, suggesting that the expression of *cis*-NAT partners could be tightly regulated through a process of competitive transcriptional interference. Under such circumstances, *cis*-NATs might be expected to exhibit reciprocal expression, which holds true for many of the antisense partners in the eukaryotic genome (http://www.narna.ncl.ac.uk).

An antisense transcript that may function as a negative regulator of gene expression by transcriptional interference has been identified in plants (Kapranov et al. 2001). In the legume *Lotus japonicus*, the expression of the late nodulin *LjNOD16* gene is controlled by a bidirectional promoter located within an intron of the gene *LjPLP-IV* (*LjPLP-IV* encodes a phosphatidylinositol transfer-like protein). Transcription from the opposite strand gives rise to an antisense transcript responsible for the control of *LjPLP-IV* expression in root nodules, where its level is significantly lower than in flowers (Kapranov et al. 2001). Similarly, during XCI, it was suggested that the *Tsix* transcripts regulate the asymmetric expression of *Xist* by an antisense mechanism (Lee et al. 1999; Sun et al. 2006). However, this mechanism of transcriptional interference cannot fully explain the repressive effect of the *Air*/*Igf2r* and *KCNQ1* loci, as genes outside of the region of overlapping antisense *Air* and *Kcnq1ot1* are also transcriptionally repressed.

RNA masking

Formation of RNA duplexes between sense and antisense transcripts might mask key regulatory features within either transcript, thereby inhibiting the interaction of important *trans*-acting factors. This form of steric inhibition could affect any step in gene expression involving protein–RNA interactions, including pre-mRNA processing, transport, translation, and degradation. An example of this method of antisense regulation is the inhibition of alternative splicing induced by the Rev-ErbAα transcript in different B-cell lines, which overlaps one of two functionally anatagonistic splice forms of the thyroid hormone receptor ErbAα2 mRNA (Hastings et al. 1997, 2000). An antisense RNA-based mechanism has also been shown to be responsible for the regulation of the human *HFE* gene, which is implicated in iron metabolism and involved in the human inherited disorder hereditary hemochromatosis (Thenie et al. 2001). Although there is no direct evidence for the role of the *HFE* antisense transcript in vivo, in vitro studies demonstrated that the antisense transcript represses the translation of the *HFE* mRNA (Thenie et al. 2001).

DsRNA-dependent mechanisms and RNAi

The interaction of antisense partners can also affect gene expression via the activation of dsRNA-dependent pathways. These might include RNA editing or RNAi-dependent gene silencing. In the first scenario, nuclear adenosine deaminases (ADARs) recognize dsRNA regions of the RNA (in such cases, the dsRNA regions generated by sense and antisense RNAs) and catalyze the hydrolytic deamination of the adenosines to inosines, A-to-I editing (Bass 2002). A small number of editing events within the coding region of the mRNA can change the coding potential of the transcript. However, long (>100 base pairs [bp]) duplexes, such as those that could result from antisense transcription, can be hyperedited such that ~50% of the adenosines on each strand are deaminated (DeCerbo and Carmichael 2005). These hyperedited RNAs may either be retained in the nucleus or degraded, thus regulating gene expression (Kumar and Carmichael 1997; Scadden and Smith 2001; Peters et al. 2003; Prasanth et al. 2005).

The formation of dsRNAs may also induce gene silencing via RNAi pathways. When small dsRNA is introduced into most eukaryotic cells, it is efficiently cleaved by the enzyme Dicer into 21- to 23-nt duplexes, termed siRNAs (Hannon 2002; Hammond 2005). These fragments then target the specific destruction of homologous mRNAs. Such sense–antisense RNA-induced gene silencing has been well documented in the case of the silencing of *Drosophila* stellate repeats (Aravin et al. 2001, 2004). Modulation of the hyperexpression of *Drosophila* stellate repeats in testis is essential for male fertility, and studies have shown that the Su(Ste) repeats produce both sense and antisense RNAs that form dsRNA in vivo (Aravin et al. 2001). The dsRNA is then cleaved to form heterogeneous 25- to 27-nt RNA species, which in turn are involved in the silencing of the stellate repeats (Aravin et al. 2001). Recent studies have also highlighted the importance of a new class of repeat-associated siRNAs (rasiRNAs) in the *Drosophila* germline

Downloaded from www.genesdev.org on September 4, 2007

(Aravin et al. 2003; Saito et al. 2006; Vagin et al. 2006). The rasiRNAs consist of 24- to 29-nt RNAs transcribed primarily from the antisense strand of repetitive sequences such as retrotransposons and heterochromatin. These RNAs associate with the P-element-induced wimpy testis (Piwi; a subclass of Argonaute) proteins, and mutations in the Piwi class of genes causes derepressed retrotransposon silencing coupled with altered levels of rasiRNAs in both the male and female germline (Vagin et al. 2006). Furthermore, in vitro studies imply that Piwi protein functions as a nuclear RNA slicer by associating specifically with rasiRNAs originating from repetitive targets (Saito et al. 2006). These results suggest that the rasiRNAs might be involved in genomic stability by silencing endogenous selfish genetic elements such as retrotransposons and repetitive sequences. Similar to rasiRNAs in flies, recent studies have also identified a similar class of germline-specific small RNAs in mammalian cells that interact with mammalian Piwi orthologs (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Lau et al. 2006; Watanabe et al. 2006). The mammalian counterpart of rasiRNAs are called piRNAs (Piwi-interacting RNAs) (Aravin et al. 2006; Girard et al. 2006; Kim 2006). Deep sequencing of piRNA sequences revealed that they correspond to regions of the genome previously thought not to be transcribed (Aravin et al. 2006; Girard et al. 2006). Unlike the rasiRNAs in *Drosophila*, the piRNA-coding regions in mammalian cells are underrepresented with repetitive sequences (Aravin et al. 2006; Carthew 2006; Girard et al. 2006). In rat testis, piRNAs form an RNP complex (piRC) that contains rat homologs of Piwi (Riwi) and RecQ1 (Lau et al. 2006). Interestingly, the piRC can cleave RNA targets in a manner dependent on piRNA complementarity, much like Ago2 cleavage of siRNA targets, suggesting the involvement of piRNAs in germline-specific transcriptional or post-transcriptional gene silencing (Carthew 2006; Lau et al. 2006).

The siRNA-induced silencing mechanism has also been proposed to explain the silencing of heterochromatin in *Schizosaccharomyces pombe* (Volpe et al. 2002). Overlapping transcription from centromeric and interspersed repetitive elements produce dsRNA, which is cleaved by the RNAi machinery and then guides recruitment of heterochromatin proteins to the repetitive elements and subsequent transcriptional silencing (Grewal and Rice 2004; Verdel and Moazed 2005). Furthermore, the purification of an RNAi effector complex in *S. pombe* termed RITS (RNA-induced initiation of transcriptional gene silencing), which is required for heterochromatin assembly, has revealed that in addition to protein components, it also contains small RNAs that are homologous to the centromeric repeats (Verdel et al. 2004). The requirement of RNA(s) in heterochromatin organization has been postulated in mammalian cells (Maison et al. 2002); however, the involvement of NATs to establish gene silencing remains to be demonstrated.

Given the diverse ways in which NATs can affect the expression of eukaryotic genes, it is hardly surprising that changes in antisense transcription can lead to abnormal patterns of gene expression that in turn contribute to pathological phenotypes. In mouse, many long antisense ncRNAs are transcribed from the complementary strand of protein-coding genes that are involved in development and disease (Furuno et al. 2006). Most of the imprinted loci contain NATs and have been suggested to play important roles in the parental-specific expression of their protein-coding partners (see "Roles of ncRNAs in Genomic Imprinting: One Is Enough" above). For example, abnormal expression of large NATs from the human *15q11–13* region has been suggested to be involved in the reduced expression of *Ube3a* that is associated with PWS/AS (Chamberlain and Brannan 2001). Another mechanism by which antisense transcription might contribute to disease is the generation of abnormal antisense transcripts that result from chromosomal rearrangements. For example, Tufarelli and coworkers (Tufarelli et al. 2003) described a novel disease mechanism leading to an inherited form of α-thalassemia in which the *hemoglobin α-2* gene is silenced by a *cis-* acting antisense RNA. An expressed sequence tag (EST)-based bioinformatics analysis suggests a possible connection between the up-regulation of antisense transcription and cancer (Shendure and Church 2002). An increased level of antisense ncRNAs from intronic regions in humans has been correlated with tumor differentiation in the case of prostate cancers (Reis et al. 2004).

RNAs as modulators of transcription and translation

Several different eukaryotic RNAs have been shown to bind and modulate the activities of proteins that impact various aspects of gene expression (Goodrich and Kugel 2006). Several of these regulatory RNAs are very abundant and thus were among the first ncRNAs to be discovered; yet their cellular roles have only recently been revealed. Some of these RNA modulators include those shown below.

pgc *RNA in* Drosophila

Germ cells retain the potential to develop into any tissue, making it critical that they be protected from inappropriate differentiation. During early embryogenesis, germ cells avoid differentiation by transiently and globally silencing mRNA transcription (Leatherman and Jongens 2003). In *Caenorhabditis elegans* early germ cells, an RNA-binding protein, PIE-1, is responsible for transcriptional inhibition by interfering with transcription elongation or associated RNA processing steps (Mello et al. 1996; Seydoux et al. 1996; Zhang et al. 2003). Similarly, in *Drosophila*, a cytoplasmic ncRNA (a major transcript of 0.7 kb and a minor transcript of 1.3 kb) *polar granule component* (*pgc*) is involved in germ cell transcriptional inhibition (Nakamura et al. 1996; Deshpande et al. 2004; Martinho et al. 2004). *pgc* RNA is localized in the germ plasm, but is not required for germ cell formation, indicating that germ cell fate can be uncoupled from transcriptional quiescence (Nakamura et

Downloaded from www.genesdev.org on September 4, 2007

al. 1996). In the absence of *pgc* RNA, pole cells express various individual genes (Blackwell 2004), indicating that transcriptional silencing is abrogated (Deshpande et al. 2004; Martinho et al. 2004). Although the exact mechanism by which *pgc* RNA attenuates transcription in germ cells is not clear, it is hypothesized that *pgc* RNA might sequester a critical transcription factor that is responsible for transcriptional elongation (Martinho et al. 2004). Alternatively, polar granules might require *pgc* RNA in order to send a signal directing early germ cells to silence transcription, or *pgc* RNA might have a fundamental role in germ cell function even though it is not required for pole cell formation (Blackwell 2004).

7SK RNA

The mammalian 7SK RNA (~330 nt) is transcribed by RNA pol III and was among the first ncRNAs to be identified (Zieve and Penman 1976). Its sequence is conserved between mouse and human (Blencowe 2002). 7SK function was revealed by the discovery that the RNA bound to and inhibited the RNA pol II transcription elongation factor P-TEFb, which consists of a CDK9-cyclin T1 or T2 or K heterodimer (Nguyen et al. 2001; Yang et al. 2001). 7SK also interacts with HEXIM1/2, which together with P-TEFb forms the inactive P-TEFb complex that cannot engage in the transcriptional elongation-dependent phosphorylation of the CTD of RNA pol II (Yik et al. 2003, 2004; Michels et al. 2004). A variety of stress conditions lead to 7SK RNA release from the protein complex, resulting in P-TEFb activation. Immunolocalization studies have revealed that a proportion of the protein components of the P-TEFb complex are localized in nuclear speckles, also known as interchromatin granule clusters (IGCs) (Haaland et al. 2005). The IGCs also contain the nuclear enriched 7SK RNA, and depleting 7SK RNA from human cells results in the mislocalization of IGC constituents (K.V. Prasanth, M. Camiolo, and D.L. Spector, unpubl.).

Mouse B2 RNA

B2 RNA (178 nt) is expressed by RNA pol III from short interspersed repetitive elements (SINEs) in the mouse genome. Expression of this RNA is increased up to 100-fold in response to environmental stresses such as heat shock (Liu et al. 1995; Allen et al. 2004). Coimmunoprecipitation and binding experiments have provided evidence that B2 RNA associates with RNA pol II upon heat shock, and in vivo and in vitro transcription experiments have revealed that B2 RNA inhibits RNA pol II by preventing the formation of active preinitiation complexes (Allen et al. 2004; Espinoza et al. 2004). The increased level of B2 RNA present in heat-shocked cells negatively regulates the expression of the genes that need to be silenced during stress, primarily by inhibiting RNA pol II initiation. Interestingly, B2 RNA-mediated transcriptional inhibition does not affect RNA pol II promoters of the heat-shock genes, indicating promoter specificity.

Heat shock RNA-1 (HSR-1)

Heat-shock transcription factor (HSF1) has an important role in the heat-shock response in vertebrate cells by inducing the expression of heat-shock proteins (HSPs) and other cryoprotective proteins (Sarge et al. 1991). HSF1 is present in unstressed cells in an inactive monomeric form and becomes activated by heat and other stress stimuli. In unstressed cells, the activity of HSF1 is negatively regulated by its interaction with certain HSPs including HSP90 (Voellmy 2004). Upon stress, HSF1 becomes trimerized, binds to the heat-shock elements present in the HSP promoters, and rescues the RNA pol II elongation complex from promoter-proximal arrest (Shopland et al. 1995; Shopland and Lis 1996). A recent study has shown that HSF1 activation by heat shock is an active process and is mediated by a ribonucleoprotein complex containing translation elongation factor eEF1A and a newly identified ncRNA, *HSR-1* (Shamovsky et al. 2006). *HSR-1* RNA is an ~600-nt poly(A)$^-$ RNA and shows a high degree of homology between human and rodents. *HSR-1* RNA is constitutively expressed, and its level seem unaffected by heat shock. However, in vivo immuoprecipitation studies revealed that the formation of the *HSR-1*/eEF1A complex is increased upon heat shock. Knockdown of *HSR-1* RNA impairs the heat-shock response in vivo, rendering cells thermosensitive and revealing the importance of *HSR-1* RNA in activation of the heat-shock response (Shamovsky et al. 2006). It has been suggested that during the initial stages of heat shock, the *HSR-1*/eEF1A complex may facilitate capture of HSF1 that is released from the HSP90 complex, and assist its assembly into trimers and/or increase the stability of HSF-1 trimers (Shamovsky et al. 2006).

Human steroid receptor RNA activator (SRA) RNA

The SRA (700–850 nt) was first identified in a screen for cofactors of the steroid hormone receptors (Lanz et al. 1999). It was isolated from mouse and human cells and shown to function as a specific coactivator of several steroid receptors. SRA RNA was found to be associated with a ribonucleoprotein complex containing the steroid receptor coactivator 1 (SRC-1), which is recruited by the steroid receptor. Interestingly, mutations within the potential ORF of *SRA* do not affect its activity, and the expression of different isoforms is cell-type-specific (Lanz et al. 1999, 2002). SRA RNA is also involved in post-translational regulation of nuclear receptor activity (Zhao et al. 2004). Recent evidence suggests that SRA is one example of a new class of RNAs that are also able to encode a peptide (Chooniedass-Kothari et al. 2004, 2006).

Dendritic BC1 RNA

BC1 (~150 nt) and BC200 (~200 nt) ncRNAs were identified as two cytoplasmic ncRNA transcripts expressed in the mouse and human nervous system, respectively (DeChiara and Brosius 1987; Martignetti and Brosius 1993; Cao et al. 2006). Interestingly, BC1 RNA is specifi-

Downloaded from www.genesdev.org on September 4, 2007

cally targeted to dendritic domains in neurons (Tiedge et al. 1991). *BC1* knockout mice show behavioral changes and lower survival rates as compared with normal controls (Lewejohann et al. 2004). BC1 RNA has been found to interact with the Fragile X mental retardation protein (FMRP). FMRP is an RNA-binding protein, and mutations associated with the absence of FMRP or altered expression of FMRP lead to fragile X syndrome (O'Donnell and Warren 2002). BC1 appears to promote the interaction between FMRP and other mRNAs that are known to interact with FMRP, possibly via base-pairing interactions, and thereby regulate the translation of these mRNAs at synapses (Zalfa et al. 2003, 2005).

Neuronal NRSE (neuron-restrictive silencer element) RNA

Another RNA that modulates transcription in neuronal cells is the NRSE RNA, a 20-nt dsRNA (Kuwabara et al. 2004; Cao et al. 2006). NRSE RNA is able to alter neuron-specific gene expression by interacting with the NRSF/REST transcriptional machinery, resulting in the transition from neural stem cells to differentiated neuronal cells (Kuwabara et al. 2004; Cao et al. 2006). In order to repress gene expression, the NRSF/REST complex recruits negative transcriptional regulators such as HDACs and methyl-binding proteins. The NRSE dsRNA is involved in the removal of those transcriptional silencers, thereby resulting in the activation of neuronal-specific genes. Moreover, NRSE dsRNA-dependent gene activation seems to require critical sequence homology between the NRSE/RE1 target and NRSE dsRNA. The function of the NRSE dsRNA clearly distinguishes it from other examples of small ncRNAs, resulting in the coining of the term "small modulatory RNAs" (smRNAs) for NRSE dsRNA-like transcripts (Kuwabara et al. 2004).

EBER RNAs

RNA-dependent protein kinase (PKR) is an interferon-induced factor of the cellular defense system against viral infection (Samuel 2001). The enzymatic activity of the protein depends on binding of RNA duplexes over a 24-bp length resulting in its autophosphorylation and dimerization. Once activated, PKR inactivates the eukaryotic initiation factor 2α (eIF2α), which results in the inactivation of the cellular translation apparatus. As a countermeasure against the action of PKR, some viruses such as Epstein-Barr virus (EBV) constitutively express two ncRNAs, EBER1 and EBER2, which are 167 nt and 172 nt long, respectively. EBER1 has been previously demonstrated to bind to PKR and has been suggested to confer resistance to Fas-mediated apoptosis in tissue culture cells by blocking PKR activity (Clarke et al. 1991; Sharp et al. 1993; Nanbo et al. 2005). EBER1 and EBER2 are the most abundant viral transcripts expressed during viral latency (~5×10^6 per cell) and are predominantly localized in the nucleus (Lerner et al. 1981; Howe and Steitz 1986; Fok et al. 2006). The EBER RNAs were also found to play a key role in the maintenance of the malignant phenotypes of Burkitt's lymphoma cells (Nanbo and Takada 2002).

The association of the above RNAs with transcription factors and possibly with DNA suggests a complexity of interactions rarely attributed to small RNAs. The role of ncRNAs as transcription modulators adds to the rapidly growing list of potential ncRNA functions.

RNAs: location, location, location

Proper RNA and protein localization is important for normal cellular function and embryonic development by regulating critical processes such as localized protein synthesis, formation of gradients of morphogens, and initiation of specific cell lineages. Some regulatory ncRNAs are known to regulate the localization of other RNAs and proteins within cells.

Xlsirt *(*Xenopus laevis *short interspersed repeat transcripts) RNA*

In amphibian oocytes, the correct localization of maternal mRNAs to the animal and vegetal regions determines normal embryonic development. In addition to mRNAs, the vegetal cortex of *Xenopus* oocytes also contains noncoding *Xlsirt* transcripts, which contain three to 13 repeats of a 79- to 81-nt element (Kloc et al. 1993; Allen et al. 2003). The *Xlsirt* RNAs are localized in the vegetal cortex at the early stages of oogenesis through the message transport organizer pathway (METRO), and one function attributed to this family of RNAs is in anchoring other RNAs to the vegetal cortex (Kloc and Etkin 1994). An intact 137-nt *cis*-acting element in the *Xlsirt* RNA is essential for its proper localization (Allen et al. 2003). The importance of *Xlsirt* RNAs has been shown for the localization of Vg1 mRNA, a member of the transforming growth factor β (TGFβ) family of developmental signaling molecules.

Hsr-ω *RNA*

The *Drosophila* genome responds to heat-shock stress with very active transcription at multiple heat-shock loci. These loci usually contain protein-coding genes. However, one heat-shock-induced locus (at position 93D of the polytene chromosomes) in *Drosophila melanogaster* does not appear to encode heat-shock proteins but produces several transcripts named *hsr*-ω RNAs. The *hsr*-ω gene is constitutively expressed in most of the tissues, but the transcript levels are rapidly increased upon various stresses (Bendena et al. 1991; Lakhotia 2003). It produces three transcripts all of which possess the same 5′ end but use different poly(A) sites. The longest transcript, hsr-ω-1 or hsr-ω-n RNA (10–15 kb long), is nuclear, unspliced, and collinear to the genomic DNA (Hogan et al. 1994). Furthermore, it is polyadenylated and has a stretch of several thousand nucleotides with

Downloaded from www.genesdev.org on September 4, 2007

short tandem repeats at its 3′ end. The second transcript, hsr-ω-2 or hsr-ω-pre-c, is also nuclear and ~1.9 kb long, and uses the proximal poly(A) site for polyadenylation. The third transcript, hsr-ω-3 or hsr-ω-c, is 1.2 kb long and represents the spliced product of hsr-ω-2, lacking the 700-bp intron (Garbe and Pardue 1986). The *hsr-ω* transcripts show poor sequence conservation among different *Drosophila* species, but show short stretches of homology (Lakhotia 2003). In contrast to hsr-ω-1 and hsr-ω-2 RNAs, hsr-ω-3 RNA is cytoplasmic and is associated with ribosomes with a short, poorly conserved ORF, present ~120 nt from the 5′ end in all *Drosophila* species (Fini et al. 1989). However, no actual peptide has been observed in vivo (Fini et al. 1989). The *hsr-ω* gene has been demonstrated to be crucial for proper development and viability of flies (Lakhotia 2003). The long *hsr-ω-n* transcripts are localized in specific subnuclear compartments, ω speckles, with various hnRNP proteins (Fig. 3; Prasanth et al. 2000). The ω speckles have been suggested to be the storage site of hnRNPs, and mutant larval cells lacking functional *hsr-ω* transcripts do not form ω speckles, resulting in a diffuse nuclear distribution of the hnRNPs (Prasanth et al. 2000). It has been suggested that the *hsr-ω* RNA plays the role of an organizer molecule by regulating the intranuclear trafficking and availability of hnRNPs (Prasanth et al. 2000; Lakhotia 2003). Interestingly, altered organization of ω speckles and conditional overexpression of *hsr-ω* also dominantly enhance the neurodegeneration caused by expression of proteins with expanded polyglutamine repeats in developing eye imaginal discs (Sengupta and Lakhotia 2006).

Satellite III (Sat III) transcripts

Similar to *hsr-ω* RNA in *Drosophila*, heat-shock treatment in human cells results in the production of ncRNAs from Sat III repeats (Jolly and Lakhotia 2006). These Sat III transcripts are transcribed predominantly from the pericentromeric regions of chromosome 9 (9q12 region), and the expression is strongly induced upon heat shock (Jolly et al. 2004; Rizzi et al. 2004). The Sat III transcripts are transcribed by RNA pol II, are polyadenylated, and vary in size (Jolly and Lakhotia 2006). These transcripts remain associated with their site of transcription (9q12) as large nuclear foci collectively called nuclear stress bodies (nSBs), where several factors including HSF1, splicing factors, and hnRNPs accumulate (Biamonti 2004; Jolly and Lakhotia 2006). Sat III transcripts have been suggested to play a role in the establishment and maintenance of a specific chromatin structure at the 9q12 pericentromeric region during stress as well as sequestering various RNA-binding proteins (Jolly and Lakhotia 2006).

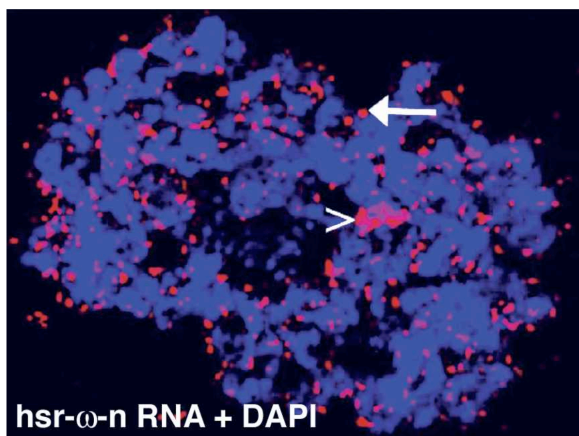

Figure 3. *Hsr-ω-n* RNA (red), visualized by RNA FISH, is localized to ω speckles in *D. melanogaster* third instar larva Malpighian tubule polytene nucleus. The arrowhead indicates the site of transcription, and the arrow shows an individual ω speckle. DNA is stained with DAPI (blue). (Image provided by Sonali Sengupta and Subhash C. Lakhotia, Cytogenetics Laboratory, Banaras Hindu University, Varanasi, India.)

NRON (noncoding repressor of NFAT [nuclear factor of activated T cells]) RNA

Shultz and colleagues (Willingham et al. 2005) recently undertook an RNAi-based genetic screen in mammalian cell lines and identified the involvement of several ncRNAs in various cellular pathways. One of the ncRNAs identified in the screen, NRON RNA (0.8–3.7 kb), acts as a repressor of NFAT. NFAT is a transcription factor responsive to local changes in calcium signals, is essential for the T-cell receptor-mediated immune response, and plays a critical role in the development of heart and vasculature, musculature, and nervous tissue (Hogan et al. 2003). Upon stimulation, the calcium-regulated phosphatase, calcineurin, dephosphorylates cytoplasmic NFAT, resulting in its nuclear translocation, thereby activating downstream pathways (Im and Rao 2004). The knockdown of NRON RNA in various cell lines resulted in significantly increased NFAT activity. NRON RNA interacts with proteins including members of the β-importin superfamily and a calmodulin-binding protein (IQGAP1), all of which show a repressive effect on NFAT activity. Further studies have revealed that NRON RNA is in a complex with members of the β-importin family and negatively regulates the nuclear trafficking of NFAT, rather than directly modulating its transcriptional activity (Willingham et al. 2005). The mode of action of NRON ncRNA has highlighted an example of how a ncRNA (NRON) inhibits the activity of a transcription factor (NFAT) primarily by preventing its nuclear import.

Pseudogene transcripts: no more 'junk'

Pseudogenes are generally considered disabled copies of functional genes that have been retained in the genome during evolution (Harrison et al. 2002; Zhang et al. 2006). Pseudogenes bear sequence similarities to a specific protein-coding gene but are unable to produce functional proteins due to the existence of frameshifts, premature stop codons, or other deleterious mutations. Most of

Downloaded from www.genesdev.org on September 4, 2007

these pseudogene sequences are the result of LINE1-mediated retrotransposition (processed pseudogene) or genome duplication (duplicated pseudogene). The human genome is estimated to contain up to 20,000 pseudogenes, and it has been predicted that ~3% of them are transcribed (Yano et al. 2004; Zhang et al. 2006). Pseudogenes are important, as they may represent genomic fossils that can be used to infer the ancestral sequence and evolutionary history of present day genes (Zhang et al. 2006) and/or they may represent a means for additional levels of gene regulation.

Recent studies have suggested the importance of pseudogene transcripts in gene regulation (Hirotsune et al. 2003; Lee 2003; Yano et al. 2004). Hirotsune et al. (2003) reported the unprecedented finding that a mouse not expressing the *Makorin1-p1* pseudogene due to a gene insertion at the *Makorin1-p1* chromosomal locus shows abnormal expression of the functional protein-coding *Makorin1* gene located elsewhere in the genome. Examination of the *Makorin1-p1* sequence indicated that it is riddled with insertions, deletions, and numerous nucleotide substitutions relative to the *Makorin1* gene. The pseudogene also has an in-frame premature stop codon. Further differences between the gene and the pseudogene include the fact that the *Makorin1-p1* mRNA contains only the first 700 nt of the *Makorin1* mRNA. Moreover, whereas both alleles of the *Makorin1* gene are transcribed, *Makorin1-p1* is paternally imprinted (Hirotsune et al. 2003). Furthermore, when the paternal *Makorin1-p1* pseudogene is disrupted, expression of *Makorin1* is markedly reduced in embryos and throughout birth. Interestingly, this mouse line died shortly after birth from multiorgan failure, suggesting an important role for *Makorin1-p1*. From these results Yano et al. (2004) have proposed that the *Makorin1-p1* RNA functions to stabilize the *Makorin1* mRNA. However, a recent study by Nicholls and colleagues (Gray et al. 2006) challenged the role of *Makorin1-p1* in regulating *Makorin1* mRNA stability. Gray et al. (2006) provided evidence that both alleles of *Makorin1-p1* in mice are hypermethylated and transcriptionally silent and therefore cannot stabilize the *Makorin1* mRNA in *trans*. Furthermore, mice in which *Makorin1* has been directly disrupted showed none of the phenotypes attributed to the partial reduction of *Makorin1* (Gray et al. 2006).

A role of pseudogene transcripts in gene regulation has also been reported in the snail *Lymnea stagnalis* (Korneev et al. 1999, 2005). Transcription of a pseudogene that is homologous to the *neuronal nitric oxide synthase* (*nNOS*) gene in a population of neurons decreases the expression levels for the *nNOS* gene. Interestingly, RNA isolated from these neuronal cells confirmed the presence of an in vivo stable RNA–RNA duplex between these two transcripts suggested to arise via a reverse-complement sequence found at the 5′ end of the *pseudo-NOS* transcript (Korneev et al. 1999). This study implicated the *pseudo-NOS* transcript as a natural antisense regulator of *nNOS* protein synthesis (Korneev et al. 1999).

The variety of known or suspected functions of pseudogene transcripts discovered to date suggests that pseudogenes as a whole have a wide range of previously unsuspected functions. In fact, a recent study has suggested that *Xist* RNA has probably evolved in eutherians through the pseudogenization of a protein-coding gene present in marsupials (Duret et al. 2006). Therefore, many pseudogenes may not represent evolutionary relics as once thought. Now more than ever, the examination of pseudogenes for unrealized functions should be evaluated in a systematic and large-scale manner.

Nuclear retained regulatory RNAs: something for a stressful day

In eukaryotic cells, protein-coding mRNAs are transported to the cytoplasm for translation. However, several earlier studies demonstrated a population of poly(A)$^+$ RNAs that were enriched in the nucleus of mammalian cells (Herman et al. 1976; Carter et al. 1991; Visa et al. 1993; Huang et al. 1994), although the identity or function(s) of these RNAs was unclear. In order to understand the roles of nuclear retained RNAs present in IGCs, a subnuclear domain that regulates pre-mRNA processing (Lamond and Spector 2003) and associated structures, RNA components of the IGCs from mouse liver nuclei were purified and RNA-FISH-based screening of mouse cell lines revealed several RNA candidates with interesting nuclear localization patterns (Prasanth et al. 2005; K.V. Prasanth and D.L. Spector, unpubl.). One such RNA identified was CTN-RNA (*Cat2-transcribed nuclear RNA*), an ~8-kb poly(A)$^+$ transcript that localizes to paraspeckles (Fig. 4; Fox et al. 2002; Prasanth et al. 2005). CTN-RNA is transcribed by the *mouse cationic amino acid transporter-2 (mCAT-2* or *Slc7a2*) gene. Other than CTN-RNA, *mCAT-2* also encodes for the protein-coding *mCAT-2* mRNA (~4.2 kb) by differential promoter and poly(A) site usage (Prasanth et al. 2005). Both *mCAT-2* mRNA and CTN-RNA are completely spliced and processed transcripts containing the same ORF, both have a unique 5′UTR, and in addition, CTN-RNA has a long unique 3′UTR and is retained in the nucleus. The 3′UTR of CTN-RNA contains inverted repeat-sequence elements of SINE origin that can form duplex RNA structures, which are extensively A-to-I edited by a member of the ADAR class of nuclear adenosine deaminases (Bass 2002; Prasanth et al. 2005). The A-to-I editing of CTN-RNA and its further interaction with a nuclear RNP complex containing p54/nrb, PSF, and matrin 3 is primarily responsible for its nuclear retention (Das and Carmichael 2005; Prasanth et al. 2005). A similar mechanism of nuclear retention of viral RNAs through A-to-I hyperediting and interaction with the cellular p54/nrb RNP complex has been previously described (Zhang and Carmichael 2001; DeCerbo and Carmichael 2005). Interestingly, knockdown of CTN-RNA, using specific antisense oligonucleotides, not only down-regulates CTN-RNA but also destabilizes *mCAT2* mRNA, suggesting a role for the nuclear CTN-RNA in stabilizing its protein-coding partner (Prasanth et al. 2005).

The cytoplasmic *mCAT-2* mRNA encodes for the

Downloaded from www.genesdev.org on September 4, 2007

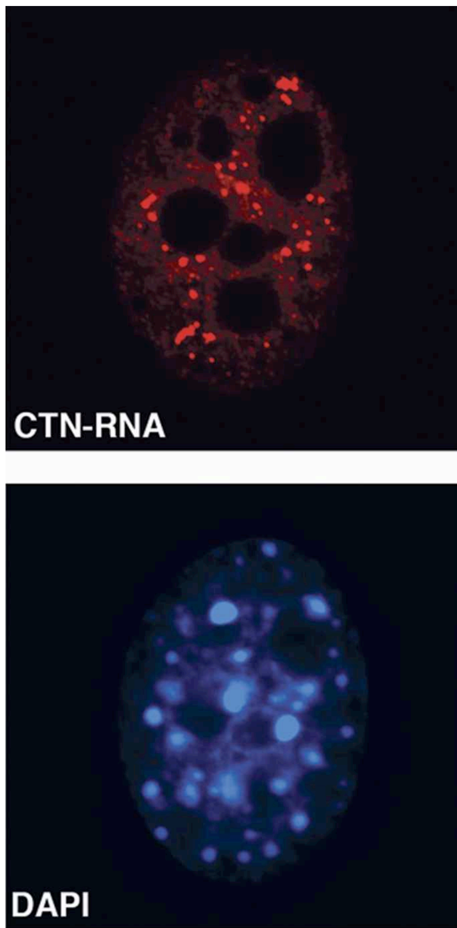

Figure 4. CTN-RNA (red), visualized by RNA FISH in mouse embryonic fibroblasts, is enriched in specific subnuclear domains, paraspeckles. DNA is stained with DAPI (blue).

CAT2 protein, which is a plasma membrane receptor that facilitates the cellular uptake of cationic amino acids, including arginine, lysine, and ornithine (MacLeod 1996). Extracellular L-arginine uptake, being the primary substrate for the synthesis of nitric oxide (NO) by *NOS* enzymes, is tightly regulated at the level of CAT2 synthesis (Lee et al. 2003). The NO pathway is induced by various stress responses and is an important component of the cellular defense program. In unstressed conditions, cells contain large amounts of nuclear CTN-RNA and basal levels of cytoplasmic *mCAT2* mRNA (Prasanth et al. 2005). Upon stress (such as IFN-γ and LPS stimulation), CTN-RNA is cleaved within its 3′UTR so as to eliminate the nuclear retention element. This cleaved mCAT2-like RNA is then exported to the cytoplasm. Due to the high levels of CTN-RNA stored in the nucleus, stress can result in the rapid cleavage of CTN-RNA and accumulation of very high levels of translation-competent mCAT2-like mRNA in the cytoplasm (Bass et al. 2005; Das and Carmichael 2005; Prasanth et al. 2005).

Quite interestingly, studies by Kay et al. (2005) have revealed a potentially similar mechanism of RNA regulation in the case of an oncofetal cytokine, *Migration-stimulating factor* (*MSF*). Similar to the *mCAT-2* locus, the *MSF* gene encoding a truncated isoform of the fibronectin gene produces two transcripts by differential poly(A) site selection (Kay et al. 2005). Both of these mRNAs have identical coding sequence and differ only in the length of their intron-derived 3′UTRs. The 5.9-kb MSF RNA was enriched in the nuclear fraction, whereas the 2.1-kb mRNA was enriched in the cytoplasmic fraction and codes for MSF. MSF-secreting fetal fibroblasts have a significantly lower nuclear level of the 5.9-kb mRNA and correspondingly higher cytoplasmic level of the 2.1-kb transcript than their nonsecreting adult counterparts. Adult fibroblasts induced to secrete MSF by treatment with transforming growth factor-β1 displayed similar changes in their respective levels of MSF mRNA. Based on this, it was suggested that expression of the MSF protein is regulated by 3′UTR cleavage of the 5.9-kb nuclear-sequestered "precursor" MSF mRNA (Kay et al. 2005).

These two examples of gene regulation have revealed an entirely new cellular mechanism for rapid post-transcriptional production of cytoplasmic mRNAs in which a protein-coding-capable RNA exists in a stable storage form in the nucleus until the cell encounters a signal that induces its post-transcriptional processing and further transport to the cytoplasm for translation. Such a role is analogous to the presence of some cytoplasmic transcription factors such as glucocorticoid receptor, which is rapidly imported into the nucleus upon cellular signals (Hager et al. 2004). The rapid response mechanism of nuclear RNA release for protein synthesis is likely to be a general paradigm for the production of numerous critical regulatory proteins.

New roles for RNAs?

RNAs as alternate genome cache

Recent studies have suggested the involvement of RNA components in rather unexpected cellular functions (Blower et al. 2005; Lolle et al. 2005; Rassoulzadegan et al. 2006). Studies by Pruitt and colleagues (Lolle et al. 2005) in *Arabidopsis thaliana* hypothesized the possibility of RNA as a messenger of non-Mendelian inheritance of extragenomic information. The study was initiated with an *A. thaliana* mutant called *hothead* (*hth*) in which various organs are fused. Several independent *hothead* mutant strains yielded apparently normal progeny at a high frequency. It was shown that this was due to precise reversion of the mutant *hothead* gene to the wild-type gene (Lolle et al. 2005). The reversion was neither due to a drastic increase in the mutation rate at the *hothead* locus nor was it due to a gene conversion, in which a related gene from elsewhere in the genome was being used as a template. Sequencing of multiple independently reverted *hth* alleles revealed no DNA sequence changes other than the restoration of the mutant nucleotide to the wild type, indicating that the process of sequence change is nonrandom. Most interestingly, the

Downloaded from www.genesdev.org on September 4, 2007

hothead mutant progeny at later generations could recover DNA variants that came from one of their great-grandparents, even if their immediate parent did not contain the variant. Lolle et al. (2005) speculated that the genetic reversion observed in the progenies is the result of a template-directed process that uses RNA as a template instead of DNA. Even though Lolle et al. (2005) failed to identify any RNA molecule that could act as a template, they speculate that this stable RNA(s), possibly in a double-stranded form, can be replicated and transmitted over multiple generations and must, under certain circumstances, be capable of modifying the DNA sequence of the nuclear genome to restore sequence information cached from previous generations. However, a recent report from Jacobson and colleagues (Peng et al. 2006) has questioned the involvement of RNA(s) in the *hth* revertants, instead arguing that increased frequency of outcrossing in *hth* mutants is responsible for the revertant phenotype. Lolle et al. (2005, 2006), however, claim that the genetic events seen in the *hth* mutants cannot be solely attributed to outcrossing. Although the identification and characterization of the RNA(s) needs to be thoroughly investigated, this entirely new concept could envisage a mechanism wherein additional allelic information is maintained in the form of RNA for several generations that could be used under conditions that compromised the continued functioning of the organism.

A recent study of the mouse *Kit* locus showed another example of RNA-mediated non-Mendelian inheritance of an epigenetic phenomenon called paramutation (Rassoulzadegan et al. 2006). Paramutation is a process in which phenotypic changes caused by an allele of a gene in one generation are remembered in subsequent generations, even if the allele that is responsible for the change is not transmitted (Soloway 2006). Heterozygous mice containing one copy of the null *Kit* allele (Kit^{tm1Alf}) and a copy of the wild-type allele exhibit white spotting on their tail tips and show reduced expression of *Kit* (Rassoulzadegan et al. 2006). Interestingly, when these heterozygous mice are crossed with wild-type mates, the wild-type progeny showed the same white spotting and reduced *Kit* mRNA levels as their heterozygous parent, even though they were fully wild-type with respect to the alleles and lacked the null allele that caused spotting in their heterozygous parent (Rassoulzadegan et al. 2006). Moreover, the wild-type progeny that exhibited the tail phenotype also had an accumulation of a mixture of smaller aberrant RNAs with sequences that matched various parts of the *Kit* mRNAs. It is possible that these RNAs are transmitted to the next generation upon fertilization, even if the allele from which they are produced is not passed on. Interestingly, when total RNA from tissues containing the aberrant *Kit* RNAs were injected into the fertilized mouse eggs, many of the progeny exhibited the "white tail spotting" phenotype (Rassoulzadegan et al. 2006). These results suggest the possibility that the aberrant RNAs arising from the Kit^{tm1Alf} alleles include some regulatory RNA(s) that cause paramutation upon transmission to the next generation. The aberrant RNAs produced from one of the alleles may regulate the corresponding wild-type allele or its transcribed mRNA, effectively silencing it (Soloway 2006). The silencing can therefore be propagated if these RNAs are packaged into germ cells and carried into the next generation. This allows successive generations to show the specific phenotype, even if the allele that caused it is not transmitted.

NcRNAs in human brain evolution

The completion of human and chimpanzee genome sequences has provided opportunities for comparative genomics toward understanding the *Homo sapiens* evolution. Previous studies to elucidate genome evolution across species mostly concentrated on processes that result in either addition or deletion of genes or on changes in amino acid sequences (Bustamante et al. 2005; Nielsen et al. 2005). However, recent studies have appreciated the importance of noncoding segments in the genome (regulatory elements, splicing signals, and ncRNA genes) in various aspects of species evolution (Amadio and Walsh 2006). Recently, Haussler and colleagues (Pollard et al. 2006) have undertaken a genome-wide scan for regions highly conserved across mammalian genomes that appear to have undergone a sudden and rapid evolution in the human lineage, and their studies suggest the involvement of ncRNA genes in shaping human brain evolution (Pollard et al. 2006). Pollard et al. (2006) classified 34,498 highly conserved regions in the genome, among which 49 "human accelerated regions (*HAR1-49*)" displayed a significantly accelerated nucleotide substitution rate only in the human genome. *HAR1*, a previously uncharacterized 118-bp region, is highly conserved across amniotes, but shows an estimated 18 fixed nucleotide substitutions in human. Interestingly, the HAR1 region is transcribed as part of two overlapping ncRNA genes, *HAR1F* and *HAR1R*, and lacks homology with any known ncRNA genes (Pollard et al. 2006). The human *HAR1F* RNA adopts a unique structural confirmation appreciably distinct from the *HAR1F* RNA of human/chimpanzee ancestors (Amadio and Walsh 2006; Pollard et al. 2006). Human embryonic brain sections showed strong expression of only *HAR1F* and not *HAR1R,* between 7 and 19 wk of gestation—a critical period for cortical neuron migration and fate specification. Moreover, *HAR1F* RNA appeared to be coexpressed with the cortical patterning protein Reelin in Cajal-Retzius neurons, a part of the cortex that is especially well developed in humans. Given the sequence overlap and tissue-regulated expression of *HAR1F* and *HAR1R* RNAs, Pollard et al. (2006) suggest the possibility of antisense regulation between these two transcripts (see Natural Antisense Transcripts [NATs]: New Players in the Gene Regulatory Network). From the sequence and expression profile analysis of the HAR1 region, Pollard et al. (2006) proposed the possible involvement of these ncRNA genes in human brain evolution. However, future studies of all *HARs*, especially *HAR1*, are needed to determine the true extent of their role in brain evolution.

Downloaded from www.genesdev.org on September 4, 2007

Regulatory RNAs implicated in complex diseases: dark side of RNA

Many of the regulatory RNAs described in humans as well as in other mammals have been linked with certain complex diseases including congenital syndromes, neurobehavioral and developmental disorders, and cancer (Table 1). The changes in expression levels or genetic and epigenetic alterations affecting ncRNAs accompanying the malignant processes strongly support a functional role of ncRNAs in normal cellular development and differentiation (Szymanski et al. 2005).

Linking ncRNA and cancer

Recent expression analysis comparing tumor cells to normal cells has shown changes in the expression of certain ncRNAs in several forms of cancer (Costa 2005; Hall and Russell 2005; Esquela-Kerscher and Slack 2006). Alterations in the methylation status of DMR, upstream of the *H19* gene, results in the loss of *H19* and/or *Igf2* biallelic expression and results in malignant cell growth (Ulaner et al. 2003). A loss of *H19* RNA has been reported in many pediatric cancers (DeBaun et al. 2002). However, there are contradicting reports on the exact role of *H19* RNA in cancer (Hao et al. 1993; Isfort et al. 1997; Ariel et al. 2000; Juan et al. 2000; Lottin et al. 2002). A recent study has suggested a role of c-Myc in the transcriptional induction of *H19* during tumorigenesis (Barsyte-Lovejoy et al. 2006). Allele-specific chromatin immunoprecipitation and expression analyses indicated that c-Myc binds and drives the expression of only the maternal *H19* allele (Barsyte-Lovejoy et al. 2006). Another imprinted gene implicated in cancer is a paternally expressed antisense RNA, *PEG8/IGF2AS*, that is transcribed from the *IGF2* locus (Szymanski et al. 2005). *IGF2AS* shows significantly elevated expression levels in Wilms' tumors and several fetal tumors but not in normal kidney tissue (Okutsu et al. 2000).

Overexpression of specific ncRNAs has also been found to be a good marker for several tumors. Colon carcinoma cells show significantly higher levels of *OCC-1* (overexpressed in colon carcinoma-1) gene transcripts. *OCC-1* RNAs show tissue-specific expression patterns and are absent or expressed at very low levels in normal mucosa (Pibouin et al. 2002). In prostatic tumors, two ncRNA genes, *DD3/PCA3 (prostate cancer antigen 3)* and *PCGEM1*, are significantly overexpressed compared with normal tissue. *DD3/PCA3* expression is restricted to the prostate and can be used for diagnosis of prostate cancer (Bussemakers et al. 1999; de Kok et al. 2002). Analysis of *PCGEM1* expression in matched normal and primary tumor specimens has revealed tumor-associated overexpression in a majority of prostate tumor specimens (Srikantan et al. 2000). Interestingly, overexpression of *PCGEM1* in cell lines correlates with increased proliferation and colony formation, which has suggested its involvement in regulation of cell growth (Petrovics et al. 2004).

In non-small-cell lung cancer (NSCLC), metastasis has been shown to be associated with increased expression of the *MALAT-1* (*metastasis associated in lung adenocarcinoma transcript 1*) gene, which encodes an 8-kb ncRNA that is conserved across several species (Ji et al. 2003). Indeed, *MALAT-1* overexpression is a prognostic parameter for poor NSCLC patient survival and can be used to identify early-stage NSCLC patients who are at risk to develop metastases. In addition to NSCLC, recent studies have also reported the overexpression of *MALAT-1* in uterine endometrial stromal sarcoma and hepatocellular carcinoma (Lin et al. 2006; Yamada et al. 2006). Another ncRNA gene, *NCRMS* (*noncoding RNA in RMS*), shows elevated expression in alveolar rhabdomyosarcoma (RMS) but not in the embryonal subtype of RMS (Chan et al. 2002). The expression of *NCRMS* may indicate a deregulation of gene expression within the large chromosomal region including several genes associated with muscle development including *myf5*, *myf6*, and growth factor *Igf2.* Abnormal patterns of *NCRMS* expression were also observed in neuroblastoma and synovial sarcoma.

HIS-1 and *BIC* ncRNA genes also have been implicated in oncogenesis and growth control, but their function in normal cells is unknown (Costa 2005). *BIC* expression was shown to be frequently associated with *c-myc* activation, being preferentially activated in metastatic tumors (Tam et al. 2002). Moreover, chick oncogenicity assays have demonstrated that *BIC* could cooperate with *c-myc* in lymphomagenesis and erythroleukemogenesis (Tam et al. 2002). Interestingly, the human *BIC* gene generates two miRNAs, one of which, miR-155, is overexpressed in patients with non-Hodgkin's lymphoma and in particular Burkitt's lymphomas (van den Berg et al. 2003; Metzler et al. 2004; Eis et al. 2005; Kluiver et al. 2005).

Several lines of evidence now indicate that the deregulation of the large class of miRNAs might underlie or be a marker for some types of cancer (McManus 2003; Esquela-Kerscher and Slack 2006). These small (~20 nt long) RNA molecules play a pivotal role as regulators in eukaryotic development (He and Hannon 2004). Their potential for regulating specific post-transcriptional regulation of gene expression combined with their small size and evolutionary conservation make them ideal candidates for agents controlling complex gene networks governing cell growth and differentiation (Bartel 2004). Altered patterns of miRNAs may therefore be responsible for changes in a cell's genetic program, which in turn results in malignant growth (McManus 2003). Altered expression of tissue-specific miRNAs has been reported in colorectal cancer cell lines and in clinical samples both of adenomatus and invasive colorectal neoplasms, in gliobastomas, in pituitary tumors, and in breast cancer (Hall and Russell 2005). In Kaposi's sarcoma induced by herpes virus, virally encoded miRNAs contribute to the oncogenicity of the virus (Cai et al. 2005). Let-7 miRNA, expression of which is greatly reduced in lung cancer, has been shown to be a negative regulator of the *RAS* oncogene (Takamizawa et al. 2004; Johnson et al. 2005). The overexpression of let-7 in tissue

Downloaded from www.genesdev.org on September 4, 2007

Table 1. *Examples of ncRNAs correlated with diseases/disorders*

NcRNAs	Disease/disorder	Reference
NcRNAs with altered expression levels in cancer		
Antisense intronic ncRNAs	Prostate cancer	Reis et al. 2004
BC1	Overexpressed in several cancers	Chen et al. 1997b
BC200	Overexpressed in breast, cervix, esophagus, lung, ovary, parotid, and tongue cancer	Chen et al. 1997a; Iacoangeli et al. 2004
BCMS	B-cell neoplasia	Wolf et al. 2001
C13orf25 (miR-17-92)	Elevated expression in lymphoma	Ota et al. 2004; L. He et al. 2005; O'Donnell et al. 2005
CMPD	Campomyelic displasia	Ninomiya et al. 1996
DD3	Overexpressed in prostate cancer	Bussemakers et al. 1999
H19	Overexpressed in liver and breast cancer	Looijenga et al. 1997; Lottin et al. 2002
HIS-1	Overexpressed in myeloid leukemia	Askew et al. 1994
HOST2	Expressed in ovarian cancer cells	Rangel et al. 2003
let-7 family miRNAs	Down-regulated in lung adenocarcinoma	Takamizawa et al. 2004; Johnson et al. 2005
MALAT-1	NSCLC, endometrial sarcoma, and hepatocellular carcinoma	Ji et al. 2003; Lin et al. 2006; Yamada et al. 2006
miR-143 and miR-145	Down-regulated in colorectal cancer	Michael et al. 2003
miR-146, miR-221, and miR-222	Elevated expression in papillary thyroid carcinoma	H. He et al. 2005
miR-155/BIC	Overexpressed in Burkitt and B-cell lymphomas; overexpressed in leukemia and breast cancer	Tam et al. 2002; Metzler et al. 2004; Eis et al. 2005; Iorio et al. 2005; Tam and Dahlberg 2006
miR-15a and miR-16-1	Deleted or down-regulated in B-cell lymphocytic leukaemia (B-CLL) and pituitary adenoma	Calin et al. 2002; Bottoni et al. 2005
miR-21	Elevated expression in glioblastoma cells and breast cancer	Chan et al. 2005; Iorio et al. 2005
miR-372 and miR-373	Testicular germ cell tumors	Voorhoeve et al. 2006
NC612	Prostate cancer	A.P. Silva et al. 2003
NCRMS	Elevated expression in alveolar rhabdomyosarcoma	Chan et al. 2002
OCC1	Overexpressed in colon carcinoma	Pibouin et al. 2002
PCGEM1	Overexpressed in prostate cancer	Srikantan et al. 2000
PEG8/IGF2AS	Fetal tumors	Okutsu et al. 2000
SRA	Steroid receptor activated RNA isoform expressed in breast cancer	Lanz et al. 1999
TRNG10	Various cancers	Roberts et al. 1998
U50HG	snoRNA host gene; located at the chromosomal breakpoint involved in human B-cell lymphoma	Tanaka et al. 2000
NcRNAs correlated with neurological diseases/disorders		
BC200	Alzheimer's	Lukiw et al. 1992
DISC2	Schizophrenia and bipolar affective disorder	Millar et al. 2000, 2004; Blackwood et al. 2001
IPW	Prader-Willi syndrome	Wevrick et al. 1994
Prion-associated RNAs	Prion pathologies	Deleault et al. 2003; Supattapone 2004
PSZA11q14	Reduced expression in brains of patients with schizophrenia	Polesskaya et al. 2003
RAY1/ST7	Autistic disorder	Vincent et al. 2002
SCA8 (KLHL1 antisense)	Spinocerebellar ataxia type 8	Nemes et al. 2000; Mutsuddi et al. 2004
UBE3A-AS	Angelman syndrome	Chamberlain and Brannan 2001
ZNF127AS	Prader-Willi syndrome	Jong et al. 1999
NcRNAs correlated with other diseases/disorders		
22k48	HIRA intronic transcript deleted in DiGeorge syndrome	Pizzuti et al. 1999
C6orf37OS	Antisense transcript from C6orf37 locus within diffuse panbronchiolitis critical region	Matsuzaka et al. 2002
COPG2IT1	Russell-Silver syndrome	Yamasaki et al. 2000
DGCR5	Disrupted in DiGeorge syndrome	Sutherland et al. 1996
H19	Beckwith-Wiedemann syndrome	Sparago et al. 2004
LIT1	Beckwith-Wiedemann syndrome	Niemitz et al. 2004
LIT1	Romano-Ward, Jervell and Lange-Nielsen syndromes	Horike et al. 2000
MESTIT 1	Russell-Silver syndrome	T. Li et al. 2002; Nakabayashi et al. 2002
PRINS	Psoriasis	Sonkoly et al. 2005

Downloaded from www.genesdev.org on September 4, 2007

culture cell lines reduces *RAS* expression, resulting in the inhibition of growth of lung cancer cells (Takamizawa et al. 2004; Johnson et al. 2005). Finally, alterations in the expression of Dicer, the enzyme responsible for processing miRNA and siRNA, have been reported in lung cancer patients and correlates with poor prognosis (Karube et al. 2005).

A detailed analysis of the distribution of miRNA genes on human chromosomes has demonstrated that the majority are located within minimal deletion, minimal amplification, or breakpoint regions linked to certain forms of cancer and that they can act both as tumor suppressors or as oncogenes (Calin et al. 2004; Calin and Croce 2006). A region of human chromosome 13q14 frequently deleted in B-cell chronic lymphocytic leukemia (B-CLL), in addition to the *BCMS* (*B-cell neoplasia-associated gene with multiple splicing*) ncRNA gene, also harbors two miRNA genes, *miR-15* and *miR-16* (Calin et al. 2002). Similarly, several types of human lymphoma are characterized by amplification of a 13q31 locus and overexpression of the *c13orf25* gene within this locus (Ota et al. 2004). *c13orf25* serves as a host gene for a cluster of seven miRNAs (miR-17-92) that show a high degree of sequence conservation with the mouse orthologs. Markedly elevated levels of pri-miR-17-92 have been found in a large percentage of lymphoma samples, and overexpression of miR-17-92 has also been shown to promote tumor development in mice (L. He et al. 2005). The expression of the *miR-17-92* cluster is regulated by c-Myc, which also induces expression of a transcription factor, E2F1, that controls the genes responsible for the G1-to-S-phase transition. Interestingly, two miRNAs (miR-17-5p and miR-20) encoded within the *mir-17-92* cluster negatively regulate E2F1 mRNA translation (O'Donnell et al. 2005). Consequently, the *mir-17-92* cluster may act physiologically to dampen the myc-E2F cycle. It has been proposed that the increased levels of miRNAs reduce the proapoptotic response to *myc* overexpression (O'Donnell et al. 2005).

Neurological diseases

Certain ncRNAs have been mapped to chromosomal regions associated with neurobehavioral diseases, including autism, bipolar affective disorder, and schizophrenia. Several schizophrenia patients carry a balanced translocation t(1:11)(q43,q14). Within the breakpoint region of chromosome 1q43, two genes called *DISC1* and *DISC2* (*disrupted in schizophrenia 1* and *2*) were found. *DISC2* produces several transcripts 2.5–5.9 kb in length without protein-coding potential. *DISC2* is transcribed from the opposite strand, and its 3′ region overlaps with the protein-coding *DISC1* gene. The *DISC1* protein is involved in intracellular transport, cell polarity, and neuronal migration, and disruption of its function may in part contribute to some neurological defects (Mehler and Mattick 2006). It has been proposed that *DISC2* ncRNA may be involved in the regulation of *DISC1* expression (Millar et al. 2000). Interestingly, family linkage investigations suggested that the *DISC1* and *DISC2* genes might also play a role in the development of both unipolar and bipolar affective disorders (Blackwood et al. 2001). The same chromosomal translocation also disrupts another ncRNA gene, *PSZA11q14* (*putative schizophrenia-associated gene from 11q14*), which shows reduced expression in schizophrenia patients (Polesskaya et al. 2003).

A t(7;13)(q31.2;q21) translocation that disrupts the complex *RAY1/ST7* locus that encodes at least 18 transcripts from both strands has been reported in an autistic patient (Vincent et al. 2002). This locus encodes for two ncRNAs in the sense orientation (*ST7OT4* and *ST7OT3*) and another two in the antisense orientation (*ST7OT1* and *ST7OT2*). The role(s) of the ncRNAs from the *RAY1/ST7* locus is not known, but the antisense transcripts have been suggested to play a role in regulating translation of the protein-coding mRNAs (Vincent et al. 2002). Interestingly, the *RAY1/ST7* locus was also described as a tumor suppressor based on the identification of several mutations in certain cases of breast and colon cancer (Zenklusen et al. 2001).

Spincocerebellar ataxia type 8

Spinocerebellar ataxia type 8 (SCA8) is unique among triplet repeat expansion-induced neurodegenerative diseases in that the gene product suggested to be involved in the pathology is a ncRNA named *SCA8* (Mutsuddi and Rebay 2005; Ranum and Cooper 2006). SCA8 is a slow, progressive form of cerebellar ataxia, characterized by gait and limb ataxia, nystagmus, and dysarthria (Ranum and Cooper 2006). *SCA8* is transcribed from the first exon of another gene, *KLHL1*, a brain-specific actin-binding protein that is transcribed in the opposite orientation (Koob et al. 1999; Nemes et al. 2000). Transgenic mice expressing the *SCA8* expansion mutation develop a progressive neurological phenotype, demonstrating that the expression of the human gene with the expansion, but not the control repeat tract, is pathogenic (Ranum and Cooper 2006). Although the exact function of *SCA8* in normal and pathological conditions is still not understood, *Drosophila* that overexpress human *SCA8* showed a late-onset progressive retina neurodegeneration (Mutsuddi et al. 2004).

Psoriasis

Recently, it has been demonstrated that overexpression of a ncRNA, *PRINS* (*Psoriasis susceptibility-related RNA gene induced by stress*), is associated with psoriasis susceptibility (Sonkoly et al. 2005). *PRINS* is transcribed by RNA pol II and is expressed at different levels in various human tissues. RNA analysis showed elevated levels of *PRINS* RNA in the uninvolved epidermis of psoriatic patients compared with both psoriatic lesions and healthy epidermis. *PRINS* expression is elevated in cells that are exposed to certain stress conditions including viral infection and ultraviolet-B irradiation (Sonkoly et al. 2005). Gene-specific silencing of *PRINS* by RNAi has

Downloaded from www.genesdev.org on September 4, 2007

revealed that down-regulation of *PRINS* impairs cell viability after serum starvation but not under normal serum conditions. These findings suggest that *PRINS* RNA functions as a regulatory ncRNA, playing a protective role in cells exposed to stress.

Summary

We have seen a recent explosion in the identification of ncRNAs, yet we have only begun to understand the complexity of these RNAs and how the cell uses these regulatory ncRNAs for various aspects of gene expression. Recent studies have revealed that some ncRNA genes, including miRNAs in human, play important roles in cell growth and differentiation, and their aberrant expression can manifest various growth abnormalities, including cancer. In addition, the expression of a variety of large and small ncRNAs has also been correlated with various neurological disorders (Table 1). Some of these ncRNAs that are implicated in various diseases have murine homologs that display similar genomic organization and expression patterns, suggesting that they may operate key regulatory networks that are conserved in eukaryotic cells.

The mechanism of action and biological roles played by the regulatory ncRNAs are extremely diverse, ranging from their involvement in dosage compensation, imprinting, and gene silencing to modulating transcription and translation. Extensive studies in the field of mammalian and *Drosophila* dosage compensation clearly demonstrate the task played by regulatory ncRNAs in X-chromosome modifications and also in regulating chromosome-specific gene expression (Heard and Disteche 2006; Spencer and Lee 2006). Quite strikingly, most of the imprinted loci in mammals are associated with antisense ncRNAs. This association suggests a possible role of ncRNAs in influencing allele-specific gene expression. There are several examples of ncRNAs that regulate gene expression by controlling the intracellular localization and stability of other RNAs and proteins (Goodrich and Kugel 2006). Furthermore, intergenic transcripts, NATs, and pseudogene transcripts with no protein-coding capacity are a widespread characteristic of eukaryotic genomes, and new functions of these transcripts are clearly emerging from recent transcriptome analysis. Thus far, we have only had a glimpse of the functions played by ncRNAs, and many surprises are likely to be revealed as further ncRNAs are identified and characterized. It has been argued that the majority of the genome in humans and other complex organisms is devoted to extensive, but hitherto largely hidden, regulatory networks that are *trans*-acted by noncoding regulatory RNAs that are essential for the evolution of complex organisms (Mattick 2001, 2003, 2004a,b,c; Mattick and Gagen 2001).

Identification and characterization of the complete compliment of ncRNAs in the genome, "RNomics," is essential before we can achieve a full understanding of the myriad of ways in which ncRNAs function in gene regulation (Huttenhofer et al. 2005). Bioinformatics is still in a relatively early stage with regard to being able to recognize functional long regulatory ncRNAs within genomic sequences (Rivas and Eddy 2001; Liu et al. 2006). Therefore, new bioinformatics tools are essential to mine the genome for regulatory ncRNAs and provide hints as to their function. Several ncRNA databases have recently been developed in order to catalog this growing class of regulatory transcripts (http://biobases.ibch.poznan.pl/ncRNA; http://www.prl.msu.edu/PLANTncRNAs/database.html; http://selab.wustl.edu/people/sls/WBhtml; http://noncode.bioinfo.org.cn; http://www.sanger.ac.uk/Software/Rfam; http://research.imb.uq.edu.au/rnadb/default.aspx).

ncRNAs are emerging as new and exciting players in gene regulatory networks, and their deregulation may underlie or be a marker for many complex diseases. Therefore, elucidating the different mechanisms of action of ncRNAs will provide not only a basic biological understanding of molecular function but will provide a critical nexus for revealing the basis of ncRNAs in disease etiology and their use as targets in subsequent drug design.

Acknowledgments

We thank Dr. John S. Mattick (Institute for Molecular Bioscience, University of Queensland, Australia), Dr. Supriya G. Prasanth, and Spector laboratory members, especially Jeremy Wilusz, Hongjae Sunwoo, and Mathew Camiolo, for critically reading the manuscript and for their input and valuable suggestions. We also thank Drs. Edith Heard, Mitzi Kuroda, and Subhash Lakhotia for generously providing images. Research in the Spector laboratory is funded by NIGMS 42694 and 71407.

References

Akhtar, A. 2003. Dosage compensation: An intertwined world of RNA and chromatin remodelling. *Curr. Opin. Genet. Dev.* **13:** 161–169.

Albrecht, U., Sutcliffe, J.S., Cattanach, B.M., Beechey, C.V., Armstrong, D., Eichele, G., and Beaudet, A.L. 1997. Imprinted expression of the murine Angelman syndrome gene, Ube3a, in hippocampal and Purkinje neurons. *Nat. Genet.* **17:** 75–78.

Allen, L., Kloc, M., and Etkin, L.D. 2003. Identification and characterization of the Xlsirt *cis*-acting RNA localization element. *Differentiation* **71:** 311–321.

Allen, T.A., Von Kaenel, S., Goodrich, J.A., and Kugel, J.F. 2004. The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nat. Struct. Mol. Biol.* **11:** 816–821.

Amadio, J.P. and Walsh, C.A. 2006. Brain evolution and uniqueness in the human genome. *Cell* **126:** 1033–1035.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11:** 1017–1027.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. 2003. The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* **5:** 337–350.

Downloaded from www.genesdev.org on September 4, 2007

Aravin, A.A., Klenov, M.S., Vagin, V.V., Bantignies, F., Cavalli, G., and Gvozdev, V.A. 2004. Dissection of a natural RNA silencing process in the *Drosophila melanogaster* germ line. *Mol. Cell. Biol.* **24:** 6742–6750.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. 2006. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442:** 203–207.

Ariel, I., Sughayer, M., Fellig, Y., Pizov, G., Ayesh, S., Podeh, D., Libdeh, B.A., Levy, C., Birman, T., Tykocinski, M.L., et al. 2000. The imprinted H19 gene is a marker of early recurrence in human bladder carcinoma. *Mol. Pathol.* **53:** 320–323.

Ashe, H.L., Monks, J., Wijgerde, M., Fraser, P., and Proudfoot, N.J. 1997. Intergenic transcription and transinduction of the human β-globin locus. *Genes & Dev.* **11:** 2494–2509.

Askew, D.S., Li, J., and Ihle, J.N. 1994. Retroviral insertions in the murine His-1 locus activate the expression of a novel RNA that lacks an extensive open reading frame. *Mol. Cell. Biol.* **14:** 1743–1751.

Babak, T., Blencowe, B.J., and Hughes, T.R. 2005. A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. *BMC Genomics* **6:** 104.

Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. 2006. Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat. Cell Biol.* **8:** 293–299.

Bae, E., Calhoun, V.C., Levine, M., Lewis, E.B., and Drewell, R.A. 2002. Characterization of the intergenic RNA profile at abdominal-A and abdominal-B in the *Drosophila* bithorax complex. *Proc. Natl. Acad. Sci.* **99:** 16847–16852.

Barsyte-Lovejoy, D., Lau, S.K., Boutros, P.C., Khosravi, F., Jurisica, I., Andrulis, I.L., Tsao, M.S., and Penn, L.Z. 2006. The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res.* **66:** 5330–5337.

Bartel, D.P. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116:** 281–297.

Bartolomei, M.S. and Tilghman, S.M. 1997. Genomic imprinting in mammals. *Annu. Rev. Genet.* **31:** 493–525.

Bass, B.L. 2002. RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* **71:** 817–846.

Bass, B.L., Hellwig, S., and Hundley, H.A. 2005. A nuclear RNA is cut out for translation. *Cell* **123:** 181–183.

Bastepe, M., Frohlich, L.F., Linglart, A., Abu-Zahra, H.S., Tojo, K., Ward, L.M., and Juppner, H. 2005. Deletion of the NESP55 differentially methylated region causes loss of maternal GNAS imprints and pseudohypoparathyroidism type Ib. *Nat. Genet.* **37:** 25–27.

Beisel, C., Imhof, A., Greene, J., Kremmer, E., and Sauer, F. 2002. Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1. *Nature* **419:** 857–862.

Bendena, W.G., Ayme-Southgate, A., Garbe, J.C., and Pardue, M.L. 1991. Expression of heat-shock locus hsr-ω in nonstressed cells during development in *Drosophila melanogaster*. *Dev. Biol.* **144:** 65–77.

Bender, W. and Fitzgerald, D.P. 2002. Transcription activates repressed domains in the *Drosophila* bithorax complex. *Development* **129:** 4923–4930.

Biamonti, G. 2004. Nuclear stress bodies: A heterochromatin affair? *Nat. Rev. Mol. Cell Biol.* **5:** 493–498.

Blackwell, T.K. 2004. Germ cells: Finding programs of mass repression. *Curr. Biol.* **14:** R229–R230.

Blackwood, D.H., Fordyce, A., Walker, M.T., St Clair, D.M., Porteous, D.J., and Muir, W.J. 2001. Schizophrenia and affective disorders—Cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: Clinical and P300 findings in a family. *Am. J. Hum. Genet.* **69:** 428–433.

Blencowe, B.J. 2002. Transcription: Surprising role for an elusive small nuclear RNA. *Curr. Biol.* **12:** R147–R149.

Blower, M.D., Nachury, M., Heald, R., and Weis, K. 2005. A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. *Cell* **121:** 223–234.

Bottoni, A., Piccin, D., Tagliati, F., Luchin, A., Zatelli, M.C., and degli Uberti, E.C. 2005. miR-15a and miR-16-1 down-regulation in pituitary adenomas. *J. Cell. Physiol.* **204:** 280–285.

Braidotti, G., Baubec, T., Pauler, F., Seidl, C., Smrzka, O., Stricker, S., Yotova, I., and Barlow, D.P. 2004. The Air noncoding RNA: An imprinted *cis*-silencing transcript. *Cold Spring Harb. Symp. Quant. Biol.* **69:** 55–66.

Brannan, C.I., Dees, E.C., Ingram, R.S., and Tilghman, S.M. 1990. The product of the H19 gene may function as an RNA. *Mol. Cell. Biol.* **10:** 28–36.

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., and Rastan, S. 1992. 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.

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J., and Willard, H.F. 1992. 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.

Bussemakers, M.J., van Bokhoven, A., Verhaegh, G.W., Smit, F.P., Karthaus, H.F., Schalken, J.A., Debruyne, F.M., Ru, N., and Isaacs, W.B. 1999. DD3: A new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res.* **59:** 5975–5979.

Bustamante, C.D., Fledel-Alon, A., Williamson, S., Nielsen, R., Hubisz, M.T., Glanowski, S., Tanenbaum, D.M., White, T.J., Sninsky, J.J., Hernandez, R.D., et al. 2005. Natural selection on protein-coding genes in the human genome. *Nature* **437:** 1153–1157.

Cai, X., Lu, S., Zhang, Z., Gonzalez, C.M., Damania, B., and Cullen, B.R. 2005. Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells. *Proc. Natl. Acad. Sci.* **102:** 5570–5575.

Calin, G.A. and Croce, C.M. 2006. Genomics of chronic lymphocytic leukemia microRNAs as new players with clinical significance. *Semin. Oncol.* **33:** 167–173.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., et al. 2002. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci.* **99:** 15524–15529.

Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., et al. 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci.* **101:** 2999–3004.

Cao, X., Yeo, G., Muotri, A.R., Kuwabara, T., and Gage, F.H. 2006. Noncoding RNAs in the mammalian central nervous system. *Annu. Rev. Neurosci.* **29:** 77–103.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. 2005. The transcriptional landscape of the mammalian genome. *Science* **309:** 1559–1563.

Carter, K.C., Taneja, K.L., and Lawrence, J.B. 1991. Discrete nuclear domains of poly(A) RNA and their relationship to

Downloaded from www.genesdev.org on September 4, 2007

the functional organization of the nucleus. *J. Cell Biol.* **115:** 1191–1202.

Carthew, R.W. 2006. Molecular biology. A new RNA dimension to genome control. *Science* **313:** 305–306.

Cassidy, S.B., Dykens, E., and Williams, C.A. 2000. Prader-Willi and Angelman syndromes: Sister imprinted disorders. *Am. J. Med. Genet.* **97:** 136–146.

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J., et al. 2004. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116:** 499–509.

Chamberlain, S.J. and Brannan, C.I. 2001. The Prader-Willi syndrome imprinting center activates the paternally expressed murine Ube3a antisense transcript but represses paternal Ube3a. *Genomics* **73:** 316–322.

Chan, A.S., Thorner, P.S., Squire, J.A., and Zielenska, M. 2002. Identification of a novel gene NCRMS on chromosome 12q21 with differential expression between rhabdomyosarcoma subtypes. *Oncogene* **21:** 3029–3037.

Chan, J.A., Krichevsky, A.M., and Kosik, K.S. 2005. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res.* **65:** 6029–6033.

Chaumeil, J., Le Baccon, P., Wutz, A., and Heard, E. 2006. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes & Dev.* **20:** 2223–2237.

Chen, W., Bocker, W., Brosius, J., and Tiedge, H. 1997a. Expression of neural BC200 RNA in human tumours. *J. Pathol.* **183:** 345–351.

Chen, W., Heierhorst, J., Brosius, J., and Tiedge, H. 1997b. Expression of neural BC1 RNA: Induction in murine tumours. *Eur. J. Cancer* **33:** 288–292.

Chen, J., Sun, M., Hurst, L.D., Carmichael, G.G., and Rowley, J.D. 2005. Genome-wide analysis of coordinate expression and evolution of human *cis*-encoded sense–antisense transcripts. *Trends Genet.* **21:** 326–329.

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

Chooniedass-Kothari, S., Emberley, E., Hamedani, M.K., Troup, S., Wang, X., Czosnek, A., Hube, F., Mutawe, M., Watson, P.H., and Leygue, E. 2004. The steroid receptor RNA activator is the first functional RNA encoding a protein. *FEBS Lett.* **566:** 43–47.

Chooniedass-Kothari, S., Hamedani, M.K., Troup, S., Hube, F., and Leygue, E. 2006. The steroid receptor RNA activator protein is expressed in breast tumor tissues. *Int. J. Cancer* **118:** 1054–1059.

Chureau, C., Prissette, M., Bourdet, A., Barbe, V., Cattolico, L., Jones, L., Eggen, A., Avner, P., and Duret, L. 2002. Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Res.* **12:** 894–908.

Clarke, P.A., Schwemmle, M., Schickinger, J., Hilse, K., and Clemens, M.J. 1991. Binding of Epstein-Barr virus small RNA EBER-1 to the double-stranded RNA-activated protein kinase DAI. *Nucleic Acids Res.* **19:** 243–248.

Cleary, M.A., van Raamsdonk, C.D., Levorse, J., Zheng, B., Bradley, A., and Tilghman, S.M. 2001. Disruption of an imprinted gene cluster by a targeted chromosomal translocation in mice. *Nat. Genet.* **29:** 78–82.

Clemson, C.M., Hall, L.L., Byron, M., McNeil, J., and Lawrence, J.B. 2006. The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. *Proc. Natl. Acad. Sci.* **103:** 7688–7693.

Costa, F.F. 2005. Non-coding RNAs: New players in eukaryotic biology. *Gene* **357:** 83–94.

Costanzi, C. and Pehrson, J.R. 1998. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393:** 599–601.

Cumberledge, S., Zaratzian, A., and Sakonju, S. 1990. Characterization of two RNAs transcribed from the *cis*-regulatory region of the abd-A domain within the *Drosophila* bithorax complex. *Proc. Natl. Acad. Sci.* **87:** 3259–3263.

Das, A.K. and Carmichael, G.G. 2005. Unleash the messenger. *Nat. Cell Biol.* **7:** 1054–1055.

de Kok, J.B., Verhaegh, G.W., Roelofs, R.W., Hessels, D., Kiemeney, L.A., Aalders, T.W., Swinkels, D.W., and Schalken, J.A. 2002. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res.* **62:** 2695–2698.

DeBaun, M.R., Niemitz, E.L., McNeil, D.E., Brandenburg, S.A., Lee, M.P., and Feinberg, A.P. 2002. Epigenetic alterations of H19 and LIT1 distinguish patients with Beckwith-Wiedemann syndrome with cancer and birth defects. *Am. J. Hum. Genet.* **70:** 604–611.

DeCerbo, J. and Carmichael, G.G. 2005. Retention and repression: Fates of hyperedited RNAs in the nucleus. *Curr. Opin. Cell Biol.* **17:** 302–308.

DeChiara, T.M. and Brosius, J. 1987. Neural BC1 RNA: cDNA clones reveal nonrepetitive sequence content. *Proc. Natl. Acad. Sci.* **84:** 2624–2628.

Deleault, N.R., Lucassen, R.W., and Supattapone, S. 2003. RNA molecules stimulate prion protein conversion. *Nature* **425:** 717–720.

Deshpande, G., Calhoun, G., and Schedl, P. 2004. Overlapping mechanisms function to establish transcriptional quiescence in the embryonic *Drosophila* germline. *Development* **131:** 1247–1257.

Drewell, R.A., Bae, E., Burr, J., and Lewis, E.B. 2002a. Transcription defines the embryonic domains of *cis*-regulatory activity at the *Drosophila* bithorax complex. *Proc. Natl. Acad. Sci.* **99:** 16853–16858.

Drewell, R.A., Goddard, C.J., Thomas, J.O., and Surani, M.A. 2002b. Methylation-dependent silencing at the H19 imprinting control region by MeCP2. *Nucleic Acids Res.* **30:** 1139–1144.

Duret, L., Chureau, C., Samain, S., Weissenbach, J., and Avner, P. 2006. The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science* **312:** 1653–1655.

Eddy, S.R. 2001. Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2:** 919–929.

Eis, P.S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M.F., Lund, E., and Dahlberg, J.E. 2005. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc. Natl. Acad. Sci.* **102:** 3627–3632.

Engel, N. and Bartolomei, M.S. 2003. Mechanisms of insulator function in gene regulation and genomic imprinting. *Int. Rev. Cytol.* **232:** 89–127.

Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M.A., Fiering, S., Enver, T., Martin, D.I., Kennedy, M., Keller, G., et al. 1998. The β-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse β-globin locus. *Mol. Cell* **2:** 447–455.

Espinoza, C.A., Allen, T.A., Hieb, A.R., Kugel, J.F., and Goodrich, J.A. 2004. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat. Struct. Mol. Biol.* **11:** 822–829.

Esquela-Kerscher, A. and Slack, F.J. 2006. Oncomirs—Micro-

Downloaded from www.genesdev.org on September 4, 2007

RNAs with a role in cancer. *Nat. Rev. Cancer* **6:** 259–269.
Feng, J., Bi, C., Clark, B.S., Mady, R., Shah, P., and Kohtz, J.D. 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & Dev.* **20:** 1470–1484.
Fini, M.E., Bendena, W.G., and Pardue, M.L. 1989. Unusual behavior of the cytoplasmic transcript of hsr ω: An abundant, stress-inducible RNA that is translated but yields no detectable protein product. *J. Cell Biol.* **108:** 2045–2057.
Fok, V., Friend, K., and Steitz, J.A. 2006. Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes nucleocytoplasmic shuttling. *J. Cell Biol.* **173:** 319–325.
Fournier, C., Goto, Y., Ballestar, E., Delaval, K., Hever, A.M., Esteller, M., and Feil, R. 2002. Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J.* **21:** 6560–6570.
Fox, A.H., Lam, Y.W., Leung, A.K., Lyon, C.E., Andersen, J., Mann, M., and Lamond, A.I. 2002. Paraspeckles: A novel nuclear domain. *Curr. Biol.* **12:** 13–25.
Frazer, K.A., Ueda, Y., Zhu, Y., Gifford, V.R., Garofalo, M.R., Mohandas, N., Martin, C.H., Palazzolo, M.J., Cheng, J.F., and Rubin, E.M. 1997. Computational and biological analysis of 680 kb of DNA sequence from the human 5q31 cytokine gene cluster region. *Genome Res.* **7:** 495–512.
Furuno, M., Pang, K.C., Ninomiya, N., Fukuda, S., Frith, M.C., Bult, C., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., et al. 2006. Clusters of internally primed transcripts reveal novel long noncoding RNAs. *PLoS Genet.* **2:** e37.
Garbe, J.C. and Pardue, M.L. 1986. Heat shock locus 93D of *Drosophila melanogaster*: A spliced RNA most strongly conserved in the intron sequence. *Proc. Natl. Acad. Sci.* **83:** 1812–1816.
Gesteland, R.F., Cech, T.R., and Atkins, J.F. 2006. *The RNA world: The nature of modern RNA suggests a prebiotic RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Ghanem, N., Jarinova, O., Amores, A., Long, Q., Hatch, G., Park, B.K., Rubenstein, J.L., and Ekker, M. 2003. Regulatory roles of conserved intergenic domains in vertebrate Dlx bigene clusters. *Genome Res.* **13:** 533–543.
Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442:** 199–202.
Goodrich, J.A. and Kugel, J.F. 2006. Non-coding-RNA regulators of RNA polymerase II transcription. *Nat. Rev. Mol. Cell Biol.* **7:** 612–616.
Graves, J.A. 1996. Mammals that break the rules: Genetics of marsupials and monotremes. *Annu. Rev. Genet.* **30:** 233–260.
Gray, T.A., Wilson, A., Fortin, P.J., and Nicholls, R.D. 2006. The putatively functional Mkrn1-p1 pseudogene is neither expressed nor imprinted, nor does it regulate its source gene in *trans*. *Proc. Natl. Acad. Sci.* **103:** 12039–12044.
Grewal, S.I. and Rice, J.C. 2004. Regulation of heterochromatin by histone methylation and small RNAs. *Curr. Opin. Cell Biol.* **16:** 230–238.
Gribnau, J., Diderich, K., Pruzina, S., Calzolari, R., and Fraser, P. 2000. Intergenic transcription and developmental remodeling of chromatin subdomains in the human β-globin locus. *Mol. Cell* **5:** 377–386.
Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. 2006. A novel class of small RNAs in mouse spermatogenic cells. *Genes & Dev.* **20:** 1709–1714.
Grun, D., Wang, Y.L., Langenberger, D., Gunsalus, K.C., and Rajewsky, N. 2005. microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS Comput. Biol.* **1:** e13.
Haaland, R.E., Herrmann, C.H., and Rice, A.P. 2005. siRNA depletion of 7SK snRNA induces apoptosis but does not affect expression of the HIV-1 LTR or P-TEFb-dependent cellular genes. *J. Cell. Physiol.* **205:** 463–470.
Hager, G.L., Nagaich, A.K., Johnson, T.A., Walker, D.A., and John, S. 2004. Dynamics of nuclear receptor movement and transcription. *Biochim. Biophys. Acta* **1677:** 46–51.
Hall, P.A. and Russell, S.H. 2005. New perspectives on neoplasia and the RNA world. *Hematol. Oncol.* **23:** 49–53.
Hammond, S.M. 2005. Dicing and slicing: The core machinery of the RNA interference pathway. *FEBS Lett.* **579:** 5822–5829.
Hannon, G.J. 2002. RNA interference. *Nature* **418:** 244–251.
Hao, Y., Crenshaw, T., Moulton, T., Newcomb, E., and Tycko, B. 1993. Tumour-suppressor activity of H19 RNA. *Nature* **365:** 764–767.
Harrison, P.M., Hegyi, H., Balasubramanian, S., Luscombe, N.M., Bertone, P., Echols, N., Johnson, T., and Gerstein, M. 2002. Molecular fossils in the human genome: Identification and analysis of the pseudogenes in chromosomes 21 and 22. *Genome Res.* **12:** 272–280.
Hastings, M.L., Milcarek, C., Martincic, K., Peterson, M.L., and Munroe, S.H. 1997. Expression of the thyroid hormone receptor gene, erbAα, in B lymphocytes: Alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. *Nucleic Acids Res.* **25:** 4296–4300.
Hastings, M.L., Ingle, H.A., Lazar, M.A., and Munroe, S.H. 2000. Post-transcriptional regulation of thyroid hormone receptor expression by *cis*-acting sequences and a naturally occurring antisense RNA. *J. Biol. Chem.* **275:** 11507–11513.
Haussecker, D. and Proudfoot, N.J. 2005. Dicer-dependent turnover of intergenic transcripts from the human β-globin gene cluster. *Mol. Cell. Biol.* **25:** 9724–9733.
He, L. and Hannon, G.J. 2004. MicroRNAs: Small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5:** 522–531.
He, H., Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., Calin, G.A., Liu, C.G., Franssila, K., Suster, S., et al. 2005. The role of microRNA genes in papillary thyroid carcinoma. *Proc. Natl. Acad. Sci.* **102:** 19075–19080.
He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J., et al. 2005. A microRNA polycistron as a potential human oncogene. *Nature* **435:** 828–833.
Heard, E. and Disteche, C.M. 2006. Dosage compensation in mammals: Fine-tuning the expression of the X chromosome. *Genes & Dev.* **20:** 1848–1867.
Herman, R.C., Williams, J.G., and Penman, S. 1976. Message and non-message sequences adjacent to poly(A) in steady state heterogeneous nuclear RNA of HeLa cells. *Cell* **7:** 429–437.
Hirotsune, S., Yoshida, N., Chen, A., Garrett, L., Sugiyama, F., Takahashi, S., Yagami, K., Wynshaw-Boris, A., and Yoshiki, A. 2003. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature* **423:** 91–96.
Hogan, N.C., Traverse, K.L., Sullivan, D.E., and Pardue, M.L. 1994. The nucleus-limited Hsr-ω-n transcript is a polyadenylated RNA with a regulated intranuclear turnover. *J. Cell Biol.* **125:** 21–30.
Hogan, P.G., Chen, L., Nardone, J., and Rao, A. 2003. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes & Dev.* **17:** 2205–2232.
Hogga, I. and Karch, F. 2002. Transcription through the iab-7

Downloaded from www.genesdev.org on September 4, 2007

cis-regulatory domain of the bithorax complex interferes with maintenance of Polycomb-mediated silencing. *Development* **129:** 4915–4922.

Horike, S., Mitsuya, K., Meguro, M., Kotobuki, N., Kashiwagi, A., Notsu, T., Schulz, T.C., Shirayoshi, Y., and Oshimura, M. 2000. Targeted disruption of the human LIT1 locus defines a putative imprinting control element playing an essential role in Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* **9:** 2075–2083.

Howe, J.G. and Steitz, J.A. 1986. Localization of Epstein-Barr virus-encoded small RNAs by in situ hybridization. *Proc. Natl. Acad. Sci.* **83:** 9006–9010.

Huang, S., Deerinck, T.J., Ellisman, M.H., and Spector, D.L. 1994. In vivo analysis of the stability and transport of nuclear poly(A)$^+$ RNA. *J. Cell Biol.* **126:** 877–899.

Huttenhofer, A., Schattner, P., and Polacek, N. 2005. Non-coding RNAs: Hope or hype? *Trends Genet.* **21:** 289–297.

Huynh, K.D. and Lee, J.T. 2003. Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426:** 857–862.

Iacoangeli, A., Lin, Y., Morley, E.J., Muslimov, I.A., Bianchi, R., Reilly, J., Weedon, J., Diallo, R., Bocker, W., and Tiedge, H. 2004. BC200 RNA in invasive and preinvasive breast cancer. *Carcinogenesis* **25:** 2125–2133.

Im, S.H. and Rao, A. 2004. Activation and deactivation of gene expression by Ca^{2+}/calcineurin–NFAT-mediated signaling. *Mol. Cells* **18:** 1–9.

Iorio, M.V., Ferracin, M., Liu, C.G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., et al. 2005. MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* **65:** 7065–7070.

Isfort, R.J., Cody, D.B., Kerckaert, G.A., Tycko, B., and LeBoeuf, R.A. 1997. Role of the H19 gene in Syrian hamster embryo cell tumorigenicity. *Mol. Carcinog.* **20:** 189–193.

Jacob, F. and Monod, J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3:** 318–356.

Ji, P., Diederichs, S., Wang, W., Boing, S., Metzger, R., Schneider, P.M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., et al. 2003. MALAT-1, a novel noncoding RNA, and thymosin β4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **22:** 8031–8041.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. 2005. RAS is regulated by the let-7 microRNA family. *Cell* **120:** 635–647.

Jolly, C. and Lakhotia, S.C. 2006. Human sat III and *Drosophila* hsrω transcripts: A common paradigm for regulation of nuclear RNA processing in stressed cells. *Nucleic Acids Res.* **34:** 5508–5514.

Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B.M., Khochbin, S., and Vourc'h, C. 2004. Stress-induced transcription of satellite III repeats. *J. Cell Biol.* **164:** 25–33.

Jones, B.K., Levorse, J.M., and Tilghman, S.M. 1998. Igf2 imprinting does not require its own DNA methylation or H19 RNA. *Genes & Dev.* **12:** 2200–2207.

Jong, M.T., Gray, T.A., Ji, Y., Glenn, C.C., Saitoh, S., Driscoll, D.J., and Nicholls, R.D. 1999. A novel imprinted gene, encoding a RING zinc-finger protein, and overlapping antisense transcript in the Prader-Willi syndrome critical region. *Hum. Mol. Genet.* **8:** 783–793.

Juan, V., Crain, C., and Wilson, C. 2000. Evidence for evolutionarily conserved secondary structure in the H19 tumor suppressor RNA. *Nucleic Acids Res.* **28:** 1221–1227.

Kaffer, C.R., Grinberg, A., and Pfeifer, K. 2001. Regulatory mechanisms at the mouse Igf2/H19 locus. *Mol. Cell. Biol.* **21:** 8189–8196.

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

Kanduri, C., Holmgren, C., Pilartz, M., Franklin, G., Kanduri, M., Liu, L., Ginjala, V., Ulleras, E., Mattsson, R., and Ohlsson, R. 2000. The 5′ flank of mouse H19 in an unusual chromatin conformation unidirectionally blocks enhancer–promoter communication. *Curr. Biol.* **10:** 449–457.

Kapranov, P., Routt, S.M., Bankaitis, V.A., de Bruijn, F.J., and Szczyglowski, K. 2001. Nodule-specific regulation of phosphatidylinositol transfer protein expression in *Lotus japonicus*. *Plant Cell* **13:** 1369–1382.

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

Kapranov, P., Drenkow, J., Cheng, J., Long, J., Helt, G., Dike, S., and Gingeras, T.R. 2005. Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res.* **15:** 987–997.

Karube, Y., Tanaka, H., Osada, H., Tomida, S., Tatematsu, Y., Yanagisawa, K., Yatabe, Y., Takamizawa, J., Miyoshi, S., Mitsudomi, T., et al. 2005. Reduced expression of Dicer associated with poor prognosis in lung cancer patients. *Cancer Sci.* **96:** 111–115.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., et al. 2005. Antisense transcription in the mammalian transcriptome. *Science* **309:** 1564–1566.

Kay, R.A., Ellis, I.R., Jones, S.J., Perrier, S., Florence, M.M., Schor, A.M., and Schor, S.L. 2005. The expression of migration stimulating factor, a potent oncofetal cytokine, is uniquely controlled by 3′-untranslated region-dependent nuclear sequestration of its precursor messenger RNA. *Cancer Res.* **65:** 10742–10749.

Kelley, R.L. 2004. Path to equality strewn with roX. *Dev. Biol.* **269:** 18–25.

Kim, V.N. 2006. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes & Dev.* **20:** 1993–1997.

Kishino, T., Lalande, M., and Wagstaff, J. 1997. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat. Genet.* **15:** 70–73.

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

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

Kloc, M. and Etkin, L.D. 1994. Delocalization of Vg1 mRNA from the vegetal cortex in *Xenopus* oocytes after destruction of Xlsirt RNA. *Science* **265:** 1101–1103.

Kloc, M., Spohr, G., and Etkin, L.D. 1993. Translocation of repetitive RNA sequences with the germ plasm in *Xenopus* oocytes. *Science* **262:** 1712–1714.

Kluiver, J., Poppema, S., de Jong, D., Blokzijl, T., Harms, G., Jacobs, S., Kroesen, B.J., and van den Berg, A. 2005. BIC and miR-155 are highly expressed in Hodgkin, primary mediastinal and diffuse large B cell lymphomas. *J. Pathol.* **207:** 243–249.

Kohtz, J.D. and Fishell, G. 2004. Developmental regulation of EVF-1, a novel non-coding RNA transcribed upstream of the

Downloaded from www.genesdev.org on September 4, 2007

mouse Dlx6 gene. *Brain Res. Gene Expr. Patterns* **4:** 407–412.

Koob, M.D., Moseley, M.L., Schut, L.J., Benzow, K.A., Bird, T.D., Day, J.W., and Ranum, L.P. 1999. An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat. Genet.* **21:** 379–384.

Korneev, S.A., Park, J.H., and O'Shea, M. 1999. Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an NOS pseudogene. *J. Neurosci.* **19:** 7711–7720.

Korneev, S.A., Straub, V., Kemenes, I., Korneeva, E.I., Ott, S.R., Benjamin, P.R., and O'Shea, M. 2005. Timed and targeted differential regulation of nitric oxide synthase (NOS) and anti-NOS genes by reward conditioning leading to long-term memory formation. *J. Neurosci.* **25:** 1188–1192.

Kumar, M. and Carmichael, G.G. 1997. Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc. Natl. Acad. Sci.* **94:** 3542–3547.

Kuwabara, T., Hsieh, J., Nakashima, K., Taira, K., and Gage, F.H. 2004. A small modulatory dsRNA specifies the fate of adult neural stem cells. *Cell* **116:** 779–793.

Lakhotia, S.C. 2003. *The non-coding, developmentally active and stress inducible hsrw gene of* Drosophila melanogaster *integrates post-transcriptional processing of other nuclear transcripts.* Kluwer Acdemic/Plenum Publishers, Georgetown, TX.

Lamond, A.I. and Spector, D.L. 2003. Nuclear speckles: A model for nuclear organelles. *Nat. Rev. Mol. Cell Biol.* **4:** 605–612.

Lanz, R.B., McKenna, N.J., Onate, S.A., Albrecht, U., Wong, J., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W. 1999. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell* **97:** 17–27.

Lanz, R.B., Razani, B., Goldberg, A.D., and O'Malley, B.W. 2002. Distinct RNA motifs are important for coactivation of steroid hormone receptors by steroid receptor RNA activator (SRA). *Proc. Natl. Acad. Sci.* **99:** 16081–16086.

Lareau, L.F., Green, R.E., Bhatnagar, R.S., and Brenner, S.E. 2004. The evolving roles of alternative splicing. *Curr. Opin. Struct. Biol.* **14:** 273–282.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. 2006. Characterization of the piRNA complex from rat testes. *Science* **313:** 363–367.

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

Leatherman, J.L. and Jongens, T.A. 2003. Transcriptional silencing and translational control: Key features of early germline development. *Bioessays* **25:** 326–335.

Lee, J.T. 2003. Molecular biology: Complicity of gene and pseudogene. *Nature* **423:** 26–28.

Lee, J.T. and Jaenisch, R. 1997. The (epi)genetic control of mammalian X-chromosome inactivation. *Curr. Opin. Genet. Dev.* **7:** 274–280.

Lee, J.T., Davidow, L.S., and Warshawsky, D. 1999. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.* **21:** 400–404.

Lee, J., Ryu, H., Ferrante, R.J., Morris Jr., S.M., and Ratan, R.R. 2003. Translational control of inducible nitric oxide synthase expression by arginine can explain the arginine paradox. *Proc. Natl. Acad. Sci.* **100:** 4843–4848.

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

Lerner, M.R., Andrews, N.C., Miller, G., and Steitz, J.A. 1981. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci.* **78:** 805–809.

Lewejohann, L., Skryabin, B.V., Sachser, N., Prehn, C., Heiduschka, P., Thanos, S., Jordan, U., Dell'Omo, G., Vyssotski, A.L., Pleskacheva, M.G., et al. 2004. Role of a neuronal small non-messenger RNA: Behavioural alterations in BC1 RNA-deleted mice. *Behav. Brain Res.* **154:** 273–289.

Lewis, E.B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276:** 565–570.

Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., Higgins, M., Feil, R., and Reik, W. 2004. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat. Genet.* **36:** 1291–1295.

Li, Q., Peterson, K.R., Fang, X., and Stamatoyannopoulos, G. 2002. Locus control regions. *Blood* **100:** 3077–3086.

Li, T., Vu, T.H., Lee, K.O., Yang, Y., Nguyen, C.V., Bui, H.Q., Zeng, Z.L., Nguyen, B.T., Hu, J.F., Murphy, S.K., et al. 2002. An imprinted PEG1/MEST antisense expressed predominantly in human testis and in mature spermatozoa. *J. Biol. Chem.* **277:** 13518–13527.

Lin, R., Maeda, S., Liu, C., Karin, M., and Edgington, T.S. 2006. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* Epub ahead of print July 31, 2006 (doi: 10.1038/sj.onc.1209846).

Lipshitz, H.D., Peattie, D.A., and Hogness, D.S. 1987. Novel transcripts from the ultrabithorax domain of the bithorax complex. *Genes & Dev.* **1:** 307–322.

Liu, W.M., Chu, W.M., Choudary, P.V., and Schmid, C.W. 1995. Cell stress and translational inhibitors transiently increase the abundance of mammalian SINE transcripts. *Nucleic Acids Res.* **23:** 1758–1765.

Liu, J., Gough, J., and Rost, B. 2006. Distinguishing protein-coding from non-coding RNAs through support vector machines. *PLoS Genet.* **2:** e29.

Lolle, S.J., Victor, J.L., Young, J.M., and Pruitt, R.E. 2005. Genome-wide non-Mendelian inheritance of extra-genomic information in *Arabidopsis*. *Nature* **434:** 505–509.

Lolle, S.J., Pruitt, R.E., Victor, J.L., and Young, J.M. 2006. Increased outcrossing in *hothead* mutants (Reply). *Nature* **443:** E8–E9.

Looijenga, L.H., Verkerk, A.J., De Groot, N., Hochberg, A.A., and Oosterhuis, J.W. 1997. H19 in normal development and neoplasia. *Mol. Reprod. Dev.* **46:** 419–439.

Lottin, S., Adriaenssens, E., Dupressoir, T., Berteaux, N., Montpellier, C., Coll, J., Dugimont, T., and Curgy, J.J. 2002. Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* **23:** 1885–1895.

Lucchesi, J.C., Kelly, W.G., and Panning, B. 2005. Chromatin remodeling in dosage compensation. *Annu. Rev. Genet.* **39:** 615–651.

Luikenhuis, S., Wutz, A., and Jaenisch, R. 2001. Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol. Cell. Biol.* **21:** 8512–8520.

Lukiw, W.J., Handley, P., Wong, L., and Crapper McLachlan, D.R. 1992. BC200 RNA in normal human neocortex, non-Alzheimer dementia (NAD), and senile dementia of the Alzheimer type (AD). *Neurochem. Res.* **17:** 591–597.

Lyon, M.F. 1961. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190:** 372–373.

MacLeod, C.L. 1996. Regulation of cationic amino acid transporter (CAT) gene expression. *Biochem. Soc. Trans.* **24:** 846–

Downloaded from www.genesdev.org on September 4, 2007

852.
Maison, C., Bailly, D., Peters, A.H., Quivy, J.P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. 2002. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* **30:** 329–334.
Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. 2004. Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303:** 666–669.
Mancini-Dinardo, D., Steele, S.J., Levorse, J.M., Ingram, R.S., and Tilghman, S.M. 2006. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes & Dev.* **20:** 1268–1282.
Marahrens, Y., Panning, B., Dausman, J., Strauss, W., and Jaenisch, R. 1997. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes & Dev.* **11:** 156–166.
Martens, J.A., Laprade, L., and Winston, F. 2004. Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature* **429:** 571–574.
Martens, J.A., Wu, P.Y., and Winston, F. 2005. Regulation of an intergenic transcript controls adjacent gene transcription in *Saccharomyces cerevisiae*. *Genes & Dev.* **19:** 2695–2704.
Martignetti, J.A. and Brosius, J. 1993. Neural BC1 RNA as an evolutionary marker: Guinea pig remains a rodent. *Proc. Natl. Acad. Sci.* **90:** 9698–9702.
Martinho, R.G., Kunwar, P.S., Casanova, J., and Lehmann, R. 2004. A noncoding RNA is required for the repression of RNApolII-dependent transcription in primordial germ cells. *Curr. Biol.* **14:** 159–165.
Masternak, K., Peyraud, N., Krawczyk, M., Barras, E., and Reith, W. 2003. Chromatin remodeling and extragenic transcription at the MHC class II locus control region. *Nat. Immunol.* **4:** 132–137.
Matsuzaka, Y., Tounai, K., Denda, A., Tomizawa, M., Makino, S., Okamoto, K., Keicho, N., Oka, A., Kulski, J.K., Tamiya, G., et al. 2002. Identification of novel candidate genes in the diffuse panbronchiolitis critical region of the class I human MHC. *Immunogenetics* **54:** 301–309.
Mattick, J.S. 1994. Introns: Evolution and function. *Curr. Opin. Genet. Dev.* **4:** 823–831.
Mattick, J.S. 2001. Non-coding RNAs: The architects of eukaryotic complexity. *EMBO Rep.* **2:** 986–991.
Mattick, J.S. 2003. Challenging the dogma: The hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays* **25:** 930–939.
Mattick, J.S. 2004a. RNA regulation: A new genetics? *Nat. Rev. Genet.* **5:** 316–323.
Mattick, J.S. 2004b. The hidden genetic program of complex organisms. *Sci. Am.* **291:** 60–67.
Mattick, J.S. 2004c. The regulatory architecture of the human genome. *Asia Pac. J. Clin. Nutr.* **13 (Suppl.):** S14.
Mattick, J.S. 2005. The functional genomics of noncoding RNA. *Science* **309:** 1527–1528.
Mattick, J.S. and Gagen, M.J. 2001. The evolution of controlled multitasked gene networks: The role of introns and other noncoding RNAs in the development of complex organisms. *Mol. Biol. Evol.* **18:** 1611–1630.
Mattick, J.S. and Makunin, I.V. 2005. Small regulatory RNAs in mammals. *Hum. Mol. Genet.* **14 (Spec. No. 1):** R121–R132.
Mattick, J.S. and Makunin, I.V. 2006. Non-coding RNA. *Hum. Mol. Genet.* **15 (Spec. No. 1):** R17–R29.
McManus, M.T. 2003. MicroRNAs and cancer. *Semin. Cancer Biol.* **13:** 253–258.
Meguro, M., Kashiwagi, A., Mitsuya, K., Nakao, M., Kondo, I., Saitoh, S., and Oshimura, M. 2001. A novel maternally expressed gene, ATP10C, encodes a putative aminophospholipid translocase associated with Angelman syndrome. *Nat. Genet.* **28:** 19–20.
Mehler, M.F. and Mattick, J.S. 2006. Non-coding RNA in the nervous system. *J. Physiol.* **575 (Pt 2):** 333–341.
Meller, V.H. 2000. Dosage compensation: Making 1X equal 2X. *Trends Cell Biol.* **10:** 54–59.
Meller, V.H. and Rattner, B.P. 2002. The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. *EMBO J.* **21:** 1084–1091.
Meller, V.H., Gordadze, P.R., Park, Y., Chu, X., Stuckenholz, C., Kelley, R.L., and Kuroda, M.I. 2000. Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in *Drosophila*. *Curr. Biol.* **10:** 136–143.
Mello, C.C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J.R. 1996. The PIE-1 protein and germline specification in *C. elegans* embryos. *Nature* **382:** 710–712.
Metzler, M., Wilda, M., Busch, K., Viehmann, S., and Borkhardt, A. 2004. High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. *Genes Chromosomes Cancer* **39:** 167–169.
Michael, M.Z., O'Connor, S.M., van Holst Pellekaan, N.G., Young, G.P., and James, R.J. 2003. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol. Cancer Res.* **1:** 882–891.
Michels, A.A., Fraldi, A., Li, Q., Adamson, T.E., Bonnet, F., Nguyen, V.T., Sedore, S.C., Price, J.P., Price, D.H., Lania, L., et al. 2004. Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.* **23:** 2608–2619.
Millar, J.K., Wilson-Annan, J.C., Anderson, S., Christie, S., Taylor, M.S., Semple, C.A., Devon, R.S., Clair, D.M., Muir, W.J., Blackwood, D.H., et al. 2000. Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* **9:** 1415–1423.
Millar, J.K., James, R., Brandon, N.J., and Thomson, P.A. 2004. DISC1 and DISC2: Discovering and dissecting molecular mechanisms underlying psychiatric illness. *Ann. Med.* **36:** 367–378.
Mitsuya, K., Meguro, M., Lee, M.P., Katoh, M., Schulz, T.C., Kugoh, H., Yoshida, M.A., Niikawa, N., Feinberg, A.P., and Oshimura, M. 1999. LIT1, an imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. *Hum. Mol. Genet.* **8:** 1209–1217.
Moore, T., Constancia, M., Zubair, M., Bailleul, B., Feil, R., Sasaki, H., and Reik, W. 1997. Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse Igf2. *Proc. Natl. Acad. Sci.* **94:** 12509–12514.
Morey, C. and Avner, P. 2004. Employment opportunities for non-coding RNAs. *FEBS Lett.* **567:** 27–34.
Mutsuddi, M. and Rebay, I. 2005. Molecular genetics of spinocerebellar ataxia type 8 (SCA8). *RNA Biol.* **2:** 49–52.
Mutsuddi, M., Marshall, C.M., Benzow, K.A., Koob, M.D., and Rebay, I. 2004. The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in *Drosophila*. *Curr. Biol.* **14:** 302–308.
Nakabayashi, K., Bentley, L., Hitchins, M.P., Mitsuya, K., Meguro, M., Minagawa, S., Bamforth, J.S., Stanier, P., Preece, M., Weksberg, R., et al. 2002. Identification and characterization of an imprinted antisense RNA (MESTIT1) in the human MEST locus on chromosome 7q32. *Hum. Mol. Genet.* **11:** 1743–1756.
Nakamura, A., Amikura, R., Mukai, M., Kobayashi, S., and

Downloaded from www.genesdev.org on September 4, 2007

Lasko, P.F. 1996. Requirement for a noncoding RNA in *Drosophila* polar granules for germ cell establishment. *Science* **274:** 2075–2079.

Nanbo, A. and Takada, K. 2002. The role of Epstein-Barr virus-encoded small RNAs (EBERs) in oncogenesis. *Rev. Med. Virol.* **12:** 321–326.

Nanbo, A., Yoshiyama, H., and Takada, K. 2005. Epstein-Barr virus-encoded poly(A)-RNA confers resistance to apoptosis mediated through Fas by blocking the PKR pathway in human epithelial intestine 407 cells. *J. Virol.* **79:** 12280–12285.

Navarro, P., Pichard, S., Ciaudo, C., Avner, P., and Rougeulle, C. 2005. Tsix transcription across the Xist gene alters chromatin conformation without affecting Xist transcription: Implications for X-chromosome inactivation. *Genes & Dev.* **19:** 1474–1484.

Nemes, J.P., Benzow, K.A., Moseley, M.L., Ranum, L.P., and Koob, M.D. 2000. The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). *Hum. Mol. Genet.* **9:** 1543–1551.

Newall, A.E., Duthie, S., Formstone, E., Nesterova, T., Alexiou, M., Johnston, C., Caparros, M.L., and Brockdorff, N. 2001. Primary non-random X inactivation associated with disruption of Xist promoter regulation. *Hum. Mol. Genet.* **10:** 581–589.

Nguyen, V.T., Kiss, T., Michels, A.A., and Bensaude, O. 2001. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414:** 322–325.

Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civello, D., White, T.J., et al. 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* **3:** e170.

Niemitz, E.L., DeBaun, M.R., Fallon, J., Murakami, K., Kugoh, H., Oshimura, M., and Feinberg, A.P. 2004. Microdeletion of LIT1 in familial Beckwith-Wiedemann syndrome. *Am. J. Hum. Genet.* **75:** 844–849.

Ninomiya, S., Isomura, M., Narahara, K., Seino, Y., and Nakamura, Y. 1996. Isolation of a testis-specific cDNA on chromosome 17q from a region adjacent to the breakpoint of t(12;17) observed in a patient with acampomelic campomelic dysplasia and sex reversal. *Hum. Mol. Genet.* **5:** 69–72.

Numata, K., Kanai, A., Saito, R., Kondo, S., Adachi, J., Wilming, L.G., Hume, D.A., Hayashizaki, Y., and Tomita, M. 2003. Identification of putative noncoding RNAs among the RIKEN mouse full-length cDNA collection. *Genome Res.* **13:** 1301–1306.

O'Donnell, W.T. and Warren, S.T. 2002. A decade of molecular studies of fragile X syndrome. *Annu. Rev. Neurosci.* **25:** 315–338.

O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. 2005. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435:** 839–843.

Ogawa, Y. and Lee, J.T. 2003. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol. Cell* **11:** 731–743.

Oh, H., Bai, X., Park, Y., Bone, J.R., and Kuroda, M.I. 2004. Targeting dosage compensation to the X chromosome of *Drosophila* males. *Cold Spring Harb. Symp. Quant. Biol.* **69:** 81–88.

Okamoto, I. and Heard, E. 2006. The dynamics of imprinted X inactivation during preimplantation development in mice. *Cytogenet. Genome Res.* **113:** 318–324.

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. 2004. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303:** 644–649.

Okamoto, I., Arnaud, D., Le Baccon, P., Otte, A.P., Disteche, C.M., Avner, P., and Heard, E. 2005. Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* **438:** 369–373.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., et al. 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420:** 563–573.

Okutsu, T., Kuroiwa, Y., Kagitani, F., Kai, M., Aisaka, K., Tsutsumi, O., Kaneko, Y., Yokomori, K., Surani, M.A., Kohda, T., et al. 2000. Expression and imprinting status of human PEG8/IGF2AS, a paternally expressed antisense transcript from the IGF2 locus, in Wilms' tumors. *J. Biochem.* **127:** 475–483.

O'Neill, M.J. 2005. The influence of non-coding RNAs on allele-specific gene expression in mammals. *Hum. Mol. Genet.* **14 (Spec. No. 1):** R113–R120.

Ota, A., Tagawa, H., Karnan, S., Tsuzuki, S., Karpas, A., Kira, S., Yoshida, Y., and Seto, M. 2004. Identification and characterization of a novel gene, C13orf25, as a target for 13q31–q32 amplification in malignant lymphoma. *Cancer Res.* **64:** 3087–3095.

Pang, K.C., Frith, M.C., and Mattick, J.S. 2006. Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends Genet.* **22:** 1–5.

Panning, B., Dausman, J., and Jaenisch, R. 1997. X chromosome inactivation is mediated by Xist RNA stabilization. *Cell* **90:** 907–916.

Pauler, F.M. and Barlow, D.P. 2006. Imprinting mechanisms—It only takes two. *Genes & Dev.* **20:** 1203–1206.

Peng, P., Chan, S.W.-L., Shaw, G.A., and Jacobsen, S.E. 2006. Increased outcrossing in *hothead* mutants. *Nature* **443:** pE8.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. 1996. Requirement for Xist in X chromosome inactivation. *Nature* **379:** 131–137.

Peters, N.T., Rohrbach, J.A., Zalewski, B.A., Byrkett, C.M., and Vaughn, J.C. 2003. RNA editing and regulation of *Drosophila* 4f-rnp expression by sas-10 antisense readthrough mRNA transcripts. *RNA* **9:** 698–710.

Petersen, C.P., Doench, J.G., Grishok, A., and Sharp, P.A. 2006. The biology of short RNAs. In *The RNA world: The nature of modern RNA suggests a prebiotic RNA world* (eds. R.F. Gesteland et al.), pp. 535–566. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Petrovics, G., Zhang, W., Makarem, M., Street, J.P., Connelly, R., Sun, L., Sesterhenn, I.A., Srikantan, V., Moul, J.W., and Srivastava, S. 2004. Elevated expression of PCGEM1, a prostate-specific gene with cell growth-promoting function, is associated with high-risk prostate cancer patients. *Oncogene* **23:** 605–611.

Pibouin, L., Villaudy, J., Ferbus, D., Muleris, M., Prosperi, M.T., Remvikos, Y., and Goubin, G. 2002. Cloning of the mRNA of overexpression in colon carcinoma-1: A sequence overexpressed in a subset of colon carcinomas. *Cancer Genet. Cytogenet.* **133:** 55–60.

Pizzuti, A., Novelli, G., Ratti, A., Amati, F., Bordoni, R., Mandich, P., Bellone, E., Conti, E., Bengala, M., Mari, A., et al. 1999. Isolation and characterization of a novel transcript embedded within HIRA, a gene deleted in DiGeorge syndrome. *Mol. Genet. Metab.* **67:** 227–235.

Plant, K.E., Routledge, S.J., and Proudfoot, N.J. 2001. Intergenic transcription in the human β-globin gene cluster. *Mol. Cell. Biol.* **21:** 6507–6514.

Plath, K., Mlynarczyk-Evans, S., Nusinow, D.A., and Panning, B. 2002. Xist RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.* **36:** 233–278.

Downloaded from www.genesdev.org on September 4, 2007

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. 2003. Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300:** 131–135.

Polesskaya, O.O., Haroutunian, V., Davis, K.L., Hernandez, I., and Sokolov, B.P. 2003. Novel putative nonprotein-coding RNA gene from 11q14 displays decreased expression in brains of patients with schizophrenia. *J. Neurosci. Res.* **74:** 111–122.

Pollard, K.S., Salama, S.R., Lambert, N., Lambot, M.A., Coppens, S., Pedersen, J.S., Katzman, S., King, B., Onodera, C., Siepel, A., et al. 2006. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **443:** 167–172.

Prasanth, K.V., Rajendra, T.K., Lal, A.K., and Lakhotia, S.C. 2000. ω speckles—A novel class of nuclear speckles containing hnRNPs associated with noncoding hsr-ω RNA in *Drosophila*. *J. Cell Sci.* **113:** 3485–3497.

Prasanth, K.V., Prasanth, S.G., Xuan, Z., Hearn, S., Freier, S.M., Bennett, C.F., Zhang, M.Q., and Spector, D.L. 2005. Regulating gene expression through RNA nuclear retention. *Cell* **123:** 249–263.

Prescott, E.M. and Proudfoot, N.J. 2002. Transcriptional collision between convergent genes in budding yeast. *Proc. Natl. Acad. Sci.* **99:** 8796–8801.

Rangel, L.B., Sherman-Baust, C.A., Wernyj, R.P., Schwartz, D.R., Cho, K.R., and Morin, P.J. 2003. Characterization of novel human ovarian cancer-specific transcripts (HOSTs) identified by serial analysis of gene expression. *Oncogene* **22:** 7225–7232.

Rank, G., Prestel, M., and Paro, R. 2002. Transcription through intergenic chromosomal memory elements of the *Drosophila* bithorax complex correlates with an epigenetic switch. *Mol. Cell. Biol.* **22:** 8026–8034.

Ranum, L.P. and Cooper, T.A. 2006. RNA-mediated neuromuscular disorders. *Annu. Rev. Neurosci.* **29:** 259–277.

Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. 2006. RNA-mediated non-Mendelian inheritance of an epigenetic change in the mouse. *Nature* **441:** 469–474.

Reik, A., Telling, A., Zitnik, G., Cimbora, D., Epner, E., and Groudine, M. 1998. The locus control region is necessary for gene expression in the human β-globin locus but not the maintenance of an open chromatin structure in erythroid cells. *Mol. Cell. Biol.* **18:** 5992–6000.

Reis, E.M., Nakaya, H.I., Louro, R., Canavez, F.C., Flatschart, A.V., Almeida, G.T., Egidio, C.M., Paquola, A.C., Machado, A.A., Festa, F., et al. 2004. Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer. *Oncogene* **23:** 6684–6692.

Richardson, B.J., Czuppon, A.B., and Sharman, G.B. 1971. Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. *Nat. New Biol.* **230:** 154–155.

Ripoche, M.A., Kress, C., Poirier, F., and Dandolo, L. 1997. Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes & Dev.* **11:** 1596–1604.

Rivas, E. and Eddy, S.R. 2001. Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics* **2:** 8.

Rizzi, N., Denegri, M., Chiodi, I., Corioni, M., Valgardsdottir, R., Cobianchi, F., Riva, S., and Biamonti, G. 2004. Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. *Mol. Biol. Cell* **15:** 543–551.

Roberts, T., Chernova, O., and Cowell, J.K. 1998. NB4S, a member of the TBC1 domain family of genes, is truncated as a result of a constitutional t(1;10)(p22;q21) chromosome translocation in a patient with stage 4S neuroblastoma. *Hum. Mol. Genet.* **7:** 1169–1178.

Rogan, D.F., Cousins, D.J., Santangelo, S., Ioannou, P.A., Antoniou, M., Lee, T.H., and Staynov, D.Z. 2004. Analysis of intergenic transcription in the human IL-4/IL-13 gene cluster. *Proc. Natl. Acad. Sci.* **101:** 2446–2451.

Ronshaugen, M., Biemar, F., Piel, J., Levine, M., and Lai, E.C. 2005. The *Drosophila* microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes & Dev.* **19:** 2947–2952.

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

Rougeulle, C., Cardoso, C., Fontes, M., Colleaux, L., and Lalande, M. 1998. An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat. Genet.* **19:** 15–16.

Runte, M., Huttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B., and Buiting, K. 2001. The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. *Hum. Mol. Genet.* **10:** 2687–2700.

Runte, M., Kroisel, P.M., Gillessen-Kaesbach, G., Varon, R., Horn, D., Cohen, M.Y., Wagstaff, J., Horsthemke, B., and Buiting, K. 2004. SNURF-SNRPN and UBE3A transcript levels in patients with Angelman syndrome. *Hum. Genet.* **114:** 553–561.

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

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes & Dev.* **20:** 2214–2222.

Samuel, C.E. 2001. Antiviral actions of interferons. *Clin. Microbiol. Rev.* **14:** 778–809.

Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. 2006. Noncoding RNAs of trithorax response elements recruit *Drosophila* Ash1 to ultrabithorax. *Science* **311:** 1118–1123.

Sanchez-Herrero, E. and Akam, M. 1989. Spatially ordered transcription of regulatory DNA in the bithorax complex of *Drosophila*. *Development* **107:** 321–329.

Sarge, K.D., Zimarino, V., Holm, K., Wu, C., and Morimoto, R.I. 1991. Cloning and characterization of two mouse heat shock factors with distinct inducible and constitutive DNA-binding ability. *Genes & Dev.* **5:** 1902–1911.

Scadden, A.D. and Smith, C.W. 2001. RNAi is antagonized by A → I hyper-editing. *EMBO Rep.* **2:** 1107–1111.

Schmitt, S. and Paro, R. 2004. Gene regulation: A reason for reading nonsense. *Nature* **429:** 510–511.

Schmitt, S., Prestel, M., and Paro, R. 2005. Intergenic transcription through a polycomb group response element counteracts silencing. *Genes & Dev.* **19:** 697–708.

Schwindinger, W.F., Francomano, C.A., and Levine, M.A. 1992. Identification of a mutation in the gene encoding the α subunit of the stimulatory G protein of adenylyl cyclase in McCune-Albright syndrome. *Proc. Natl. Acad. Sci.* **89:** 5152–5156.

Seidl, C.I., Stricker, S.H., and Barlow, D.P. 2006. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J.* **25:** 3565–3575.

Sengupta, S. and Lakhotia, S.C. 2006. Altered expression of the noncoding *hsrw* gene enhances poly-Q induced neurotoxic-

Downloaded from www.genesdev.org on September 4, 2007

ity in *Drosophila. RNA Biol.* **3:** 28–35.

Seydoux, G., Mello, C.C., Pettitt, J., Wood, W.B., Priess, J.R., and Fire, A. 1996. Repression of gene expression in the embryonic germ lineage of *C. elegans*. *Nature* **382:** 713–716.

Shabalina, S.A. and Spiridonov, N.A. 2004. The mammalian transcriptome and the function of non-coding DNA sequences. *Genome Biol.* **5:** 105.

Shamovsky, I., Ivannikov, M., Kandel, E.S., Gershon, D., and Nudler, E. 2006. RNA-mediated response to heat shock in mammalian cells. *Nature* **440:** 556–560.

Sharman, G.B. 1971. Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature* **230:** 231–232.

Sharp, T.V., Schwemmle, M., Jeffrey, I., Laing, K., Mellor, H., Proud, C.G., Hilse, K., and Clemens, M.J. 1993. Comparative analysis of the regulation of the interferon-inducible protein kinase PKR by Epstein-Barr virus RNAs EBER-1 and EBER-2 and adenovirus VAI RNA. *Nucleic Acids Res.* **21:** 4483–4490.

Sheardown, S.A., Duthie, S.M., Johnston, C.M., Newall, A.E., Formstone, E.J., Arkell, R.M., Nesterova, T.B., Alghisi, G.C., Rastan, S., and Brockdorff, N. 1997. Stabilization of Xist RNA mediates initiation of X chromosome inactivation. *Cell* **91:** 99–107.

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

Shibata, S. and Lee, J.T. 2003. Characterization and quantitation of differential Tsix transcripts: Implications for Tsix function. *Hum. Mol. Genet.* **12:** 125–136.

Shibata, S. and Lee, J.T. 2004. Tsix transcription- versus RNA-based mechanisms in Xist repression and epigenetic choice. *Curr. Biol.* **14:** 1747–1754.

Shopland, L.S. and Lis, J.T. 1996. HSF recruitment and loss at most *Drosophila* heat shock loci is coordinated and depends on proximal promoter sequences. *Chromosoma* **105:** 158–171.

Shopland, L.S., Hirayoshi, K., Fernandes, M., and Lis, J.T. 1995. HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites. *Genes & Dev.* **9:** 2756–2769.

Silva, A.P., Salim, A.C., Bulgarelli, A., de Souza, J.E., Osorio, E., Caballero, O.L., Iseli, C., Stevenson, B.J., Jongeneel, C.V., de Souza, S.J., et al. 2003. Identification of 9 novel transcripts and two RGSL genes within the hereditary prostate cancer region (HPC1) at 1q25. *Gene* **310:** 49–57.

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H., Jenuwein, T., Otte, A.P., and Brockdorff, N. 2003. Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 Polycomb group complexes. *Dev. Cell* **4:** 481–495.

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

Sleutels, F., Tjon, G., Ludwig, T., and Barlow, D.P. 2003. Imprinted silencing of Slc22a2 and Slc22a3 does not need transcriptional overlap between Igf2r and Air. *EMBO J.* **22:** 3696–3704.

Smith, C.A., Katz, M., and Sinclair, A.H. 2003. DMRT1 is upregulated in the gonads during female-to-male sex reversal in ZW chicken embryos. *Biol. Reprod.* **68:** 560–570.

Soloway, P.D. 2006. Genetics: Paramutable possibilities. *Nature* **441:** 413–414.

Sonkoly, E., Bata-Csorgo, Z., Pivarcsi, A., Polyanka, H., Kenderessy-Szabo, A., Molnar, G., Szentpali, K., Bari, L., Megyeri, K., Mandi, Y., et al. 2005. Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. *J. Biol. Chem.* **280:** 24159–24167.

Sparago, A., Cerrato, F., Vernucci, M., Ferrero, G.B., Silengo, M.C., and Riccio, A. 2004. Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome. *Nat. Genet.* **36:** 958–960.

Spencer, R.J. and Lee, J.T. 2006. Large noncoding RNAs in mammalian gene dosage regulation. In *The RNA world: The nature of modern RNA suggests a prebiotic RNA world* (eds. R.F. Gesteland et al.), pp. 595–630. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Srikantan, V., Zou, Z., Petrovics, G., Xu, L., Augustus, M., Davis, L., Livezey, J.R., Connell, T., Sesterhenn, I.A., Yoshino, K., et al. 2000. PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc. Natl. Acad. Sci.* **97:** 12216–12221.

Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. 2003. Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1:** E60.

Stavropoulos, N., Lu, N., and Lee, J.T. 2001. A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. *Proc. Natl. Acad. Sci.* **98:** 10232–10237.

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

Supattapone, S. 2004. Prion protein conversion in vitro. *J. Mol. Med.* **82:** 348–356.

Sutherland, H.F., Wadey, R., McKie, J.M., Taylor, C., Atif, U., Johnstone, K.A., Halford, S., Kim, U.J., Goodship, J., Baldini, A., et al. 1996. Identification of a novel transcript disrupted by a balanced translocation associated with DiGeorge syndrome. *Am. J. Hum. Genet.* **59:** 23–31.

Szymanski, M. and Barciszewski, J. 2003. Regulation by RNA. *Int. Rev. Cytol.* **231:** 197–258.

Szymanski, M. and Barciszewski, J. 2006. Regulatory RNAs in mammals. *Handb. Exp. Pharmacol.* **173:** 45–72.

Szymanski, M., Barciszewska, M.Z., Erdmann, V.A., and Barciszewski, J. 2005. A new frontier for molecular medicine: Noncoding RNAs. *Biochim. Biophys. Acta* **1756:** 65–75.

Takagi, N. and Sasaki, M. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256:** 640–642.

Takagi, N., Sugawara, O., and Sasaki, M. 1982. Regional and temporal changes in the pattern of X-chromosome replication during the early post-implantation development of the female mouse. *Chromosoma* **85:** 275–286.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., et al. 2004. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res.* **64:** 3753–3756.

Takemoto, N., Kamogawa, Y., Jun Lee, H., Kurata, H., Arai, K.I., O'Garra, A., Arai, N., and Miyatake, S. 2000. Cutting edge: Chromatin remodeling at the IL-4/IL-13 intergenic regulatory region for Th2-specific cytokine gene cluster. *J. Immunol.* **165:** 6687–6691.

Tam, W. and Dahlberg, J.E. 2006. miR-155/BIC as an oncogenic microRNA. *Genes Chromosomes Cancer* **45:** 211–212.

Tam, W., Hughes, S.H., Hayward, W.S., and Besmer, P. 2002. Avian bic, a gene isolated from a common retroviral site in avian leukosis virus-induced lymphomas that encodes a noncoding RNA, cooperates with c-myc in lymphomagenesis and erythroleukemogenesis. *J. Virol.* **76:** 4275–4286.

Downloaded from www.genesdev.org on September 4, 2007

Tanaka, R., Satoh, H., Moriyama, M., Satoh, K., Morishita, Y., Yoshida, S., Watanabe, T., Nakamura, Y., and Mori, S. 2000. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. *Genes Cells* **5:** 277–287.

Teranishi, M., Shimada, Y., Hori, T., Nakabayashi, O., Kikuchi, T., Macleod, T., Pym, R., Sheldon, B., Solovei, I., Macgregor, H., et al. 2001. Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosome Res.* **9:** 147–165.

Thakur, N., Tiwari, V.K., Thomassin, H., Pandey, R.R., Kanduri, M., Gondor, A., Grange, T., Ohlsson, R., and Kanduri, C. 2004. An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. *Mol. Cell. Biol.* **24:** 7855–7862.

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

Tiedge, H., Fremeau Jr., R.T., Weinstock, P.H., Arancio, O., and Brosius, J. 1991. Dendritic location of neural BC1 RNA. *Proc. Natl. Acad. Sci.* **88:** 2093–2097.

Tuan, D., Kong, S., and Hu, K. 1992. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc. Natl. Acad. Sci.* **89:** 11219–11223.

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

Ulaner, G.A., Vu, T.H., Li, T., Hu, J.F., Yao, X.M., Yang, Y., Gorlick, R., Meyers, P., Healey, J., Ladanyi, M., et al. 2003. Loss of imprinting of IGF2 and H19 in osteosarcoma is accompanied by reciprocal methylation changes of a CTCF-binding site. *Hum. Mol. Genet.* **12:** 535–549.

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

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313:** 320–324.

van den Berg, A., Kroesen, B.J., Kooistra, K., de Jong, D., Briggs, J., Blokzijl, T., Jacobs, S., Kluiver, J., Diepstra, A., Maggio, E., et al. 2003. High expression of B-cell receptor inducible gene BIC in all subtypes of Hodgkin lymphoma. *Genes Chromosomes Cancer* **37:** 20–28.

Verdel, A. and Moazed, D. 2005. RNAi-directed assembly of heterochromatin in fission yeast. *FEBS Lett.* **579:** 5872–5878.

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

Verona, R.I., Mann, M.R., and Bartolomei, M.S. 2003. Genomic imprinting: Intricacies of epigenetic regulation in clusters. *Annu. Rev. Cell Dev. Biol.* **19:** 237–259.

Vincent, J.B., Petek, E., Thevarkunnel, S., Kolozsvari, D., Cheung, J., Patel, M., and Scherer, S.W. 2002. The RAY1/ST7 tumor-suppressor locus on chromosome 7q31 represents a complex multi-transcript system. *Genomics* **80:** 283–294.

Visa, N., Puvion-Dutilleul, F., Harper, F., Bachellerie, J.P., and Puvion, E. 1993. Intranuclear distribution of poly(A) RNA determined by electron microscope in situ hybridization. *Exp. Cell Res.* **208:** 19–34.

Voellmy, R. 2004. On mechanisms that control heat shock transcription factor activity in metazoan cells. *Cell Stress Chaperones* **9:** 122–133.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297:** 1833–1837.

Voorhoeve, P.M., le Sage, C., Schrier, M., Gillis, A.J., Stoop, H., Nagel, R., Liu, Y.P., van Duijse, J., Drost, J., Griekspoor, A., et al. 2006. A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell* **124:** 1169–1181.

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

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. 2006. 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.

Webber, A.L., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. 1998. Location of enhancers is essential for the imprinting of H19 and Igf2 genes. *Nature* **391:** 711–715.

Wevrick, R., Kerns, J.A., and Francke, U. 1994. Identification of a novel paternally expressed gene in the Prader-Willi syndrome region. *Hum. Mol. Genet.* **3:** 1877–1882.

Williams, C.A., Angelman, H., Clayton-Smith, J., Driscoll, D.J., Hendrickson, J.E., Knoll, J.H., Magenis, R.E., Schinzel, A., Wagstaff, J., Whidden, E.M., et al. 1995. Angelman syndrome: Consensus for diagnostic criteria. Angelman Syndrome Foundation. *Am. J. Med. Genet.* **56:** 237–238.

Willingham, A.T. and Gingeras, T.R. 2006. TUF love for 'junk' DNA. *Cell* **125:** 1215–1220.

Willingham, A.T., Orth, A.P., Batalov, S., Peters, E.C., Wen, B.G., Aza-Blanc, P., Hogenesch, J.B., and Schultz, P.G. 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **309:** 1570–1573.

Wolf, S., Mertens, D., Schaffner, C., Korz, C., Dohner, H., Stilgenbauer, S., and Lichter, P. 2001. B-cell neoplasia associated gene with multiple splicing (BCMS): The candidate B-CLL gene on 13q14 comprises more than 560 kb covering all critical regions. *Hum. Mol. Genet.* **10:** 1275–1285.

Wroe, S.F., Kelsey, G., Skinner, J.A., Bodle, D., Ball, S.T., Beechey, C.V., Peters, J., and Williamson, C.M. 2000. An imprinted transcript, antisense to Nesp, adds complexity to the cluster of imprinted genes at the mouse Gnas locus. *Proc. Natl. Acad. Sci.* **97:** 3342–3346.

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

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

Wutz, A., Rasmussen, T.P., and Jaenisch, R. 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30:** 167–174.

Xu, N., Tsai, C.L., and Lee, J.T. 2006. Transient homologous chromosome pairing marks the onset of X inactivation. *Science* **311:** 1149–1152.

Yamada, K., Kano, J., Tsunoda, H., Yoshikawa, H., Okubo, C., Ishiyama, T., and Noguchi, M. 2006. Phenotypic characterization of endometrial stromal sarcoma of the uterus. *Can-*

Downloaded from www.genesdev.org on September 4, 2007

cer Sci. **97:** 106–112.

Yamasaki, K., Hayashida, S., Miura, K., Masuzaki, H., Ishimaru, T., Niikawa, N., and Kishino, T. 2000. The novel gene, γ2-COP (COPG2), in the 7q32 imprinted domain escapes genomic imprinting. *Genomics* **68:** 330–335.

Yamashita, M., Ukai-Tadenuma, M., Kimura, M., Omori, M., Inami, M., Taniguchi, M., and Nakayama, T. 2002. Identification of a conserved GATA3 response element upstream proximal from the interleukin-13 gene locus. *J. Biol. Chem.* **277:** 42399–42408.

Yang, X.J. 2005. Multisite protein modification and intramolecular signaling. *Oncogene* **24:** 1653–1662.

Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. 2001. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414:** 317–322.

Yano, Y., Saito, R., Yoshida, N., Yoshiki, A., Wynshaw-Boris, A., Tomita, M., and Hirotsune, S. 2004. A new role for expressed pseudogenes as ncRNA: Regulation of mRNA stability of its homologous coding gene. *J. Mol. Med.* **82:** 414–422.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., et al. 2003. Widespread occurrence of antisense transcription in the human genome. *Nat. Biotechnol.* **21:** 379–386.

Yik, J.H., Chen, R., Nishimura, R., Jennings, J.L., Link, A.J., and Zhou, Q. 2003. Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol. Cell* **12:** 971–982.

Yik, J.H., Chen, R., Pezda, A.C., Samford, C.S., and Zhou, Q. 2004. A human immunodeficiency virus type 1 Tat-like arginine-rich RNA-binding domain is essential for HEXIM1 to inhibit RNA polymerase II transcription through 7SK snRNA-mediated inactivation of P-TEFb. *Mol. Cell. Biol.* **24:** 5094–5105.

Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., Oostra, B., and Bagni, C. 2003. The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell* **112:** 317–327.

Zalfa, F., Adinolfi, S., Napoli, I., Kuhn-Holsken, E., Urlaub, H., Achsel, T., Pastore, A., and Bagni, C. 2005. Fragile X mental retardation protein (FMRP) binds specifically to the brain cytoplasmic RNAs BC1/BC200 via a novel RNA-binding motif. *J. Biol. Chem.* **280:** 33403–33410.

Zamore, P.D. and Haley, B. 2005. Ribo-gnome: The big world of small RNAs. *Science* **309:** 1519–1524.

Zenklusen, J.C., Conti, C.J., and Green, E.D. 2001. Mutational and functional analyses reveal that ST7 is a highly conserved tumor-suppressor gene on human chromosome 7q31. *Nat. Genet.* **27:** 392–398.

Zerucha, T., Stuhmer, T., Hatch, G., Park, B.K., Long, Q., Yu, G., Gambarotta, A., Schultz, J.R., Rubenstein, J.L., and Ekker, M. 2000. A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. *J. Neurosci.* **20:** 709–721.

Zhang, Z. and Carmichael, G.G. 2001. The fate of dsRNA in the nucleus: A p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106:** 465–475.

Zhang, F., Barboric, M., Blackwell, T.K., and Peterlin, B.M. 2003. A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. *Genes & Dev.* **17:** 748–758.

Zhang, Z., Carriero, N., Zheng, D., Karro, J., Harrison, P.M., and Gerstein, M. 2006. PseudoPipe: An automated pseudogene identification pipeline. *Bioinformatics* **22:** 1437–1439.

Zhao, X., Patton, J.R., Davis, S.L., Florence, B., Ames, S.J., and Spanjaard, R.A. 2004. Regulation of nuclear receptor activity by a pseudouridine synthase through posttranscriptional modification of steroid receptor RNA activator. *Mol. Cell* **15:** 549–558.

Zieve, G. and Penman, S. 1976. Small RNA species of the HeLa cell: Metabolism and subcellular localization. *Cell* **8:** 19–31.

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