

Pervasive transcription constitutes a new level of eukaryotic genome regulation

Julia Berretta & Antonin Morillon⁺

Centre de Génétique Moléculaire—Centre National de la Recherche Scientifique (CGM—CNRS), Université of Pierre et Marie Curie, Gif-sur-Yvette, Paris, France

During the past few years, it has become increasingly evident that the expression of eukaryotic genomes is far more complex than had been previously noted. The idea that the transcriptome is derived exclusively from protein-coding genes and some specific non-coding RNAs—such as snRNAs, snoRNAs, tRNAs or rRNAs—has been swept away by numerous studies indicating that RNA polymerase II can be found at almost any genomic location. Pervasive transcription is widespread and, far from being a futile process, has a crucial role in controlling gene expression and genomic plasticity. Here, we review recent findings that point to cryptic transcription as a fundamental component of the regulation of eukaryotic genomes.

Keywords: cryptic; non-coding; RNA; genome

EMBO reports (2009) 10, 973–982. doi:10.1038/embor.2009.181

Pervasive transcription in eukaryotic genomes

The high complexity of eukaryotic genome expression has recently started coming to light. The development of high-resolution tiling arrays, the emergence of new technologies in the field of sequencing (RNA-Seq), and large-scale chromatin immunoprecipitation experiments (ChIP-chip), in addition to cDNA-library sequencing and SAGE, has allowed us to discriminate quantitatively most of the cellular transcripts. It has also provided us with new and exciting information about the occupancy of the transcription machinery throughout the genome (Table 1). RNAPII has been found at unexpected sites—such as intergenic regions or heterochromatin domains in *Saccharomyces cerevisiae* (Steinmetz *et al*, 2006)—and at previously unannotated regions of the human genome (Kim *et al*, 2005). The genomes of diverse organisms, including yeast (David *et al*, 2006; Dutrow *et al*, 2008; Nagalakshmi *et al*, 2008; Wilhelm *et al*, 2008), plants (L. Li *et al*, 2006), *Drosophila* (Stolc *et al*, 2004) and mammals (Bertone *et al*, 2004; Carninci *et al*, 2005; Cheng *et al*, 2005; He *et al*, 2008; Kapranov *et al*, 2007a; Kim *et al*, 2005), undergo heavy transcription. For example, up to 85% of the genome is transcribed in

S. cerevisiae (David *et al*, 2006). Consistent with RNAPII localization, a large number of the transcripts arise from intronic and intergenic regions (Cheng *et al*, 2005; Dutrow *et al*, 2008; L. Li *et al*, 2006; Nagalakshmi *et al*, 2008; Stolc *et al*, 2004), raising the questions of how and from which promoter they are expressed. Another important pool of non-coding RNAs (ncRNAs) aligns with known ORFs, either in the same orientation (sense) or in opposite orientation (antisense) to the coding transcript (Carninci *et al*, 2006; He *et al*, 2008; Katayama *et al*, 2005). As much as 11% of the human transcriptome derives from this latter category (He *et al*, 2008).

Cryptic unstable and stable transcripts

A microarray analysis, based on previous SAGE data obtained from *S. cerevisiae* (Velculescu *et al*, 1997), showed that the expression of many transcripts that are encoded in intergenic regions is increased in a mutant that lacks Rrp6, a catalytic subunit of the nuclear exosome (Wyers *et al*, 2005). Due to their rapid turnover in wild-type cells, these ncRNAs—which are 200–800 nt long, are transcribed by RNAPII, capped and polyadenylated—were named cryptic unstable transcripts (CUTs). The termination of CUT transcription is dependent on the Nrd1–Nab3 pathway (Arigo *et al*, 2006; Thiebaut *et al*, 2006). Nrd1 and Nab3 are thought to recruit the TRAMP polyadenylation complex to the CUT (Thiebaut *et al*, 2006). The subsequent TRAMP-dependent polyadenylation of the CUT is believed to facilitate its degradation by targeting the exosome to the transcript (LaCava *et al*, 2005; Vanacova *et al*, 2005; Wyers *et al*, 2005). In addition to being substrates for the nuclear exosome, an abundant set of intergenic transcripts were recently shown to undergo 5' to 3' cytoplasmic decay, mainly through Xrn1 (Lee *et al*, 2008; Thompson & Parker, 2007). The Steinmetz and Jacquier laboratories have further enlightened the field by constructing the first high-resolution transcriptome architecture maps for CUTs (Neil *et al*, 2009; Xu *et al*, 2009) in *S. cerevisiae*, thereby providing additional evidence of the pervasive nature of eukaryotic transcription. By using a tiling microarray approach and comparing the transcriptomes of wild-type yeast growing under diverse conditions with those of mutants lacking Rrp6, Steinmetz and co-workers identified 7,272 transcripts. In addition to identifying ORFs (71%) and other known RNAs—such as tRNAs and snoRNAs—they found that CUTs account for 13% of these transcripts (Xu *et al*, 2009).

Centre de Génétique Moléculaire—Centre National de la Recherche Scientifique (CGM—CNRS), Université of Pierre et Marie Curie, 91198 Gif-sur-Yvette, Paris 6, France
⁺Corresponding author. Tel: +33 1 6982 3638; Fax: +33 1 6982 3877;
E-mail: morillon@cgm.cnrs-gif.fr

Submitted 28 May 2009; accepted 7 July 2009; published online 14 August 2009

Table 1 | High-throughput technology used for transcriptome analyses

Technique	Principle	Aim	Organism	Main references
SAGE	Reverse transcription to cDNA of the transcriptome Short sequence tag production from unique poly(A)+ transcripts Cloning and sequencing of sequence tag arrays Quantitation of tag recovery	Quantitative gene expression data without prerequisite of sequence knowledge	<i>S. cerevisiae</i>	Velculescu <i>et al</i> , 1997
3'-long SAGE	Similar to SAGE Introduction of a recognition site for a IIS restriction enzyme at the 3' end of cDNA, recovery of 3' end tags High-throughput sequencing	Identification of 3' ends Distinction between overlapping transcripts	<i>S. cerevisiae</i>	Neil <i>et al</i> , 2009
CAGE	Similar to SAGE but from CAP-selected RNAs	Identification of transcription start sites	Mouse Human	Carninci <i>et al</i> , 2006
ASSAGE	Similar to SAGE RNA treatment with bisulphite (changes cytosine residues to uracil) before reverse transcription	Unambiguous strand specificity	Human	He <i>et al</i> , 2008
RNA-Seq	Reverse transcription to cDNA of poly(A)+ RNA High-throughput sequencing	Quantitative transcriptome data High sensitivity, low background, high resolution	<i>S. cerevisiae</i> <i>S. pombe</i>	Nagalakshimi <i>et al</i> , 2008; Wilhelm <i>et al</i> , 2008
GRO-seq	Extension of nascent RNA associated with engaged RNAPII (nuclear run-on) Isolation and purification of the RNAs Reverse transcription to cDNA High-throughput sequencing	Genome-wide positional mapping, determination of transcript amount, orientation of engaged RNAPII	Human	Core <i>et al</i> , 2008
High-resolution tiling arrays	Reverse transcription of total or selected RNA followed by hybridization on DNA chip Probes overlap and cover the entire genome	Quantitative transcriptome data from sequenced genomes Strand specific	<i>S. cerevisiae</i> <i>S. pombe</i> Rice Human	David <i>et al</i> , 2006; Xu <i>et al</i> , 2009; Dutrow <i>et al</i> , 2008; Wilhelm <i>et al</i> , 2008; L. Li <i>et al</i> , 2006; Kapranov <i>et al</i> , 2007a; Preker <i>et al</i> , 2008
ChIP-chip	Chromatin immunoprecipitation of RNAPII and TFs Hybridization on DNA array	Mapping of chromatin-bound RNAPII and TFs, genome-wide	<i>S. cerevisiae</i> Human	Steinmetz <i>et al</i> , 2006; Kim <i>et al</i> , 2005
ChIP-Seq	Chromatin immunoprecipitation of RNAPII High-throughput sequencing	Mapping of chromatin-bound RNAPII, also at low complexity sequences (such as telomeres and transposons)	Mouse	Seila <i>et al</i> , 2008

The techniques used in the ENCODE and FANTOM projects are not detailed. RNAPII, RNA polymerase II; *S. cerevisiae*, *Saccharomyces cerevisiae*; *S. pombe*, *Schizosaccharomyces pombe*; TE, transcription factor.

Jacquier's group used a 3'-long SAGE approach followed by deep sequencing to draw a genomic map of CUTs at the nucleotide resolution. By comparing wild-type and CUT-enriched RNA fractions, they identified 1,496 CUT clusters that did not correspond to any annotated feature (Neil *et al*, 2009). The data from these two studies partly overlap; however, discrepancies in the quantity and the quality of the detected CUTs were probably observed owing to the use of different techniques and also to the different purification methods and deletion strains used.

In addition, the Steinmetz group identified a new class of transcripts that do not correspond to any previously annotated genomic feature and are detectable in wild-type yeast strains, which were appropriately named stable unannotated transcripts (SUTs). These SUTs, of unknown function at present, account for 12% of the transcripts identified by the tiling microarray (Xu *et al*, 2009).

Importantly, it is becoming clear that unstable transcripts are not unique to *S. cerevisiae*. The depletion of nuclear exosome subunits in human cells also leads to the accumulation of a new class of short, polyadenylated and highly unstable transcripts known as PROMPTs (for promoter upstream transcripts; Preker *et al*, 2008). These data raise the possibility that CUTs, and possibly SUTs, might be conserved throughout the eukaryotic kingdom.

Non-coding transcription: where does it start?

The genomic organization of unannotated transcripts (SUTs and CUTs) in *S. cerevisiae* suggests that their transcription is not random, but rather clustered in defined transcription units (Neil *et al*, 2009; Xu *et al*, 2009). The transcription start sites of 68% of the unambiguously characterized, unannotated transcripts correlate strongly with the 5' nucleosome-depleted region (5'NDR;

also known as the nucleosome-free region) of another transcriptional unit (Mavrich *et al*, 2008; Xu *et al*, 2009). Remarkably, these analyses revealed that a small fraction—less than 5%—of these promoter-associated ncRNAs (PARs) are transcribed in the same orientation as the mRNA, leading to the production of multiple, partly overlapping transcripts at several genes (Fig 1; Xu *et al*, 2009). Conversely, more than 95% of the transcripts are in antisense orientation with respect to the downstream gene and generate divergent transcripts (Xu *et al*, 2009; Fig 1). However, these percentages should be interpreted with caution, as many unresolved issues—such as the exact positioning of the nucleosomes and the precise 5' and 3' ends of all unannotated transcripts in *S. cerevisiae*—require further investigation (L. Steinmetz, personal communication). Notably, the transcript pairs that are divergently expressed have a coordinated transcription, suggesting that they result from a single bidirectional promoter (Neil *et al*, 2009; Xu *et al*, 2009). Although the number and types of identified unannotated transcript do not perfectly overlap between the Xu and Neil studies, the bidirectional character of most of the promoters that generate PARs and the large number of initiation events from shared NDRs are consistent in both studies and are strongly supported by the data. The 3'-long SAGE analysis by Neil and colleagues validates the observation that only a minority of PARs are transcribed in the same orientation as the gene, as this technique unambiguously identifies the 3' ends of the transcripts.

Polyadenylated ncRNAs also arise from non-random genomic regions in higher eukaryotes. A tiling array of the human transcriptome revealed that ~1.1% of the genome is covered with unannotated RNAs, which are smaller than 200 nt and cluster at the 5' and 3' ends of protein-coding genes (Kapranov *et al*, 2007a). In addition, three groups have recently reported the existence of diffuse divergent transcription from human and murine genomes, suggesting that promoter bidirectionality is a common and conserved feature of eukaryotic transcription. Global run-on sequencing indicates that divergent and engaged RNAPII is present at 77% of active genes in human cells (Core *et al*, 2008). Similarly, the analysis of a cDNA library from mice revealed the presence of short RNAs—between 20 and 90 nt long—located near the transcription start sites of more than 50% of protein-coding genes, and frequently arising from divergent transcription (Seila *et al*, 2008). Positionally conserved transcription initiation RNAs (tiRNAs) of fewer than 22 nt have also been shown recently to map to the transcription start sites of several protein-coding genes in humans, chicken and *Drosophila*, possibly as a result of abortive transcription events by stalling polymerases or owing to polymerase 'backtracking' events (Taft *et al*, 2009). Finally, the PROMPTs are also bidirectionally transcribed from a 1.5-kb region upstream from the transcription start sites of active protein-coding genes and depend on the same promoter as the downstream mRNA (Preker *et al*, 2008). Therefore, it seems that a large proportion of eukaryotic promoters are bidirectional, which raises questions about how they are regulated. The answer might come from yeast, in which PARs and mRNAs seem to originate from different pre-initiation complexes and might compete for the same pool of transcription factors (Neil *et al*, 2009; Fig 2A).

However, pervasive transcription not only originates from promoter regions, but also a number of ncRNAs seem to start at the 3' ends of genes, in both yeast and humans (Kapranov *et al*, 2007a; Xu *et al*, 2009; Fig 1). A 3' NDR was recently found near the mRNA

Glossary

<i>Air</i>	antisense <i>Igf2r</i> RNA
CBP	CREB binding protein
<i>CCND1</i>	cyclin D1
CpG	cytidine-phosphate-guanosine
<i>DHFR</i>	dehydrofolate reductase
<i>fbp</i>	fructose-1,6-biphosphatase
HDAC	histone deacetylase complex
HMT	histone methyl transferase
<i>HOTAIR</i>	HOX antisense intergenic RNA
<i>HOX</i>	homeobox
<i>ICL2</i>	isocytate lyase 2
<i>Igf2r</i>	insulin-like growth factor 2 receptor
<i>IMD2</i>	inosine monophosphate dehydrogenase 2
<i>Isw2</i>	imitation switch 2
<i>Kcnq1ot1</i>	KCNQ1 overlapping transcript 1
<i>KCS1</i>	pKC1 suppressor
<i>MPH2</i>	maltose permease homologue 2
<i>Nab3</i>	nuclear polyadenylated RNA-binding 3
nt	nucleotide
<i>Nrd1</i>	nuclear pre-mRNA downregulation 1
ORF	open reading frame
<i>PHO5/84</i>	phosphate metabolism 5/84
piRNA	PIWI-associated RNA
PIWI	P-element-induced wimpy testis
PRC2	polycomb repressive complex 2
<i>Rat1</i>	ribonucleic acid trafficking 1
rDNA	ribosomal DNA
RNAPII	RNA polymerase II
<i>Rpd3S</i>	reduced potassium dependency 3 small
<i>Rrp6</i>	ribosomal RNA processing 6
SAGE	Serial Analysis of Gene Expression
SET	suppressor of variegation, enhancer of zeste, trithorax group regulator
siRNA	small interfering RNA
snoRNA	small nucleolar RNA
<i>Spt6/16</i>	suppressor of Ty1 6/16
<i>SRG1</i>	SER3 regulatory gene
<i>SUC2</i>	sucrose fermentation 2
TFIIB	transcription factor II B
<i>TLS</i>	translocated in liposarcoma
TRAMP	Trf4–Air1/2–Mtr4 polyadenylation
<i>TRF4</i>	topoisomerase I requiring function 4
tRNA	transfer RNA
TY1	transposon in yeast 1
UAS	upstream activating sequence
<i>URA2</i>	uracil requiring 2
<i>Xist</i>	X (inactive)-specific transcript
<i>Xrn1</i>	5'–3' exoribonuclease 1

cleavage and polyadenylation sites, and is suspected to promote non-coding, antisense transcription (Mavrich *et al*, 2008) that is repressed by the chromatin-remodelling factor *Isw2* (Whitehouse *et al*, 2007). In the absence of *Isw2*, the nucleosomes are shifted away from the 3' intergenic region, thereby allowing the generation of cryptic antisense RNAs. In yeast, 32% of unannotated transcripts initiate at these 3' NDR, and most are transcribed in an antisense orientation with respect to the mRNA transcription (Xu *et al*, 2009).

Exon-originated cryptic transcripts have also been identified. Histone chaperones and histone modification factors—such as *Spt6*, *Spt16* and the HMT *Set2*—are known to be involved in

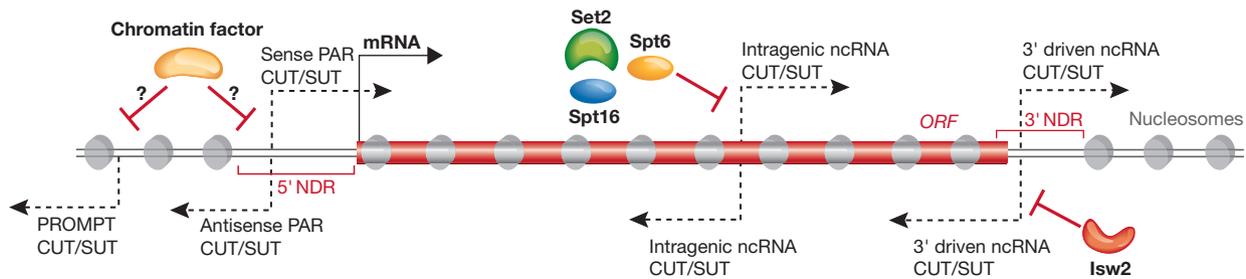


Fig 1 | Relative orientation of non-coding RNA and mRNA transcription. PARs (CUTs and SUTs) and PROMPTs can be transcribed from the gene promoter region—from the 5' NDR in particular—and from intergenic regions in either sense or antisense orientation. In yeast, transcription from 3' NDRs is mostly repressed by Isw2, and intragenic cryptic promoters are generally inhibited by Spt6, Spt16 or Set2. 5' NDR, 5' nucleosome-depleted region; CUT, cryptic unstable transcript; Isw2, imitation switch 2; ncRNA, non-coding RNA; PAR, promoter-associated ncRNA; PROMPT, promoter upstream transcript; Set2, SET-domain-containing 2; Spt6/16, suppressor of Ty1 6/16.

the repression of exon-originated cryptic transcription in yeast (Fig1). It is believed that on RNAPII passage, the evicted histones are replaced by Spt6 and Spt16 to suppress spurious transcription from hidden promoters embedded in ORFs (Carrozza *et al*, 2005; Kaplan *et al*, 2003). However, it is noteworthy that chromatin modifications control cryptic promoter activity but not the stability of the arising transcripts.

Strikingly, in contrast to the intragenic or 3' NDR cryptic transcripts, no chromatin-remodelling complex or histone modification has been shown to control the expression of PARs or PROMPTs (Fig 1). Further investigation will be necessary to determine the mechanism that controls the promoter activity of these transcripts.

In addition to single-copy genes, non-coding transcripts originate in large numbers from repetitive regions of the genome. In flies, vertebrates and plants, a particular type of short RNA arises from the transcription of retrotransposon clusters (Aravin *et al*, 2007, 2008; Brennecke *et al*, 2007). These 21–28 nt-long piRNAs are probably matured from a transposon-derived long RNA precursor (Brennecke *et al*, 2007). Similarly, in *S. cerevisiae*, the high-copy-number Ty1 retrotransposon gives rise to antisense ncRNAs that are stabilized when 5'–3' cytoplasmic RNA decay is impaired (Berretta *et al*, 2008). Surprisingly, non-coding transcripts have been found to originate from heterochromatic regions as well; RNAPII-dependent telomeric transcripts have recently been described in humans, mouse, zebrafish and yeast. The length of the telomeric transcripts ranges from 100 nt to 10 kb and their synthesis is strand-specific. Yeast telomeric RNAs are mainly unstable, and have been shown to be degraded by the 5'–3' nuclear exonuclease Rat1 (Luke *et al*, 2008), although *TRF4*-deleted strains (Houseley *et al*, 2007; M. Kwapisz and A. Morillon, unpublished data) and *xrn1Δ* mutants (M. Kwapisz and A. Morillon, unpublished data) also display strong accumulation of other species of subtelomeric RNAs, suggesting a role for these enzymes in their stability. Similarly, heterochromatic transcripts arise from the silent rDNA loci in *S. cerevisiae*. The rDNA repeats include intergenic spacers, which are refractory to transcription by RNAPII and controlled by silencing factors. In strains that lack the silencing factor, RNAPII-dependent spurious transcripts are expressed (C. Li *et al*, 2006). In addition, rDNA-originated CUTs are stabilized in strains mutated for *Nrd1*, *Trf4* or the exosome (Houseley *et al*, 2007; Vasiljeva *et al*, 2008), as occurs with CUTs generated by single-copy genes.

Spurious transcription: a trial and error signature?

The recent discovery of widespread transcription over all eukaryotic genomes that have been analysed raises the question of the role that non-coding and/or cryptic transcripts might have in the regulation of genomic plasticity. Most of the recently identified RNAs do not have protein-coding potential and could potentially be a 'side effect' of intrinsic chromatin characteristics; perhaps the mere depletion of nucleosomes allows RNAPII to transcribe. This hypothesis is supported by the strong increase in transcripts that arise from 'exonic' promoters in *Spt6*, *Spt16* or *Set2* deletion mutants, as well as by the rise of antisense transcription starting at the 3' end of protein-coding genes when *Isw2* is absent. So far, no physiological condition has been described in which hidden promoters engage in transcription, suggesting that these RNAs do not have any function. Conversely, a large number of spurious transcripts are unstable in normal growth conditions, as is exemplified by the rapid turnover of PROMPTs and CUTs. Thus, a current view is that eukaryotic transcription might act as a 'trial and error test', in which RNAPII sometimes initiates and possibly elongates non-functional transcripts, before transcribing proper protein-coding genes.

However, some features of non-coding transcription seem to contradict the idea that widespread genome expression is futile. Not all spurious transcripts are unstable: SUTs are present in wild-type yeast, although at low levels, and PARs are stable in human and mouse. In addition, several genomic sequences that encode ncRNAs are much more conserved than would be expected from a random drifting DNA region (Carninci *et al*, 2005; Kapranov *et al*, 2007b; Stolc *et al*, 2004; Xu *et al*, 2009). For example, more than 95% of the recently identified long non-coding intervening RNAs (lincRNAs) show clear evolutionary conservation (Guttman *et al*, 2009). Some transcripts have been shown to be regulated under different growth conditions (Dutrow *et al*, 2008; Xu *et al*, 2009) or during development (Stolc *et al*, 2004), suggesting that they are biologically significant and could have a potential regulatory role. In addition, a number of lincRNAs are regulated by known transcription factors (Guttman *et al*, 2009). Finally, the relative expression of promoter-associated ncRNAs is far from random: most of the yeast tandem CUT–mRNA transcript pairs—as well as the sense–antisense transcript pairs—seem to be anti-regulated (Neil *et al*, 2009; Xu *et al*, 2009), and divergent bidirectional-promoter-derived

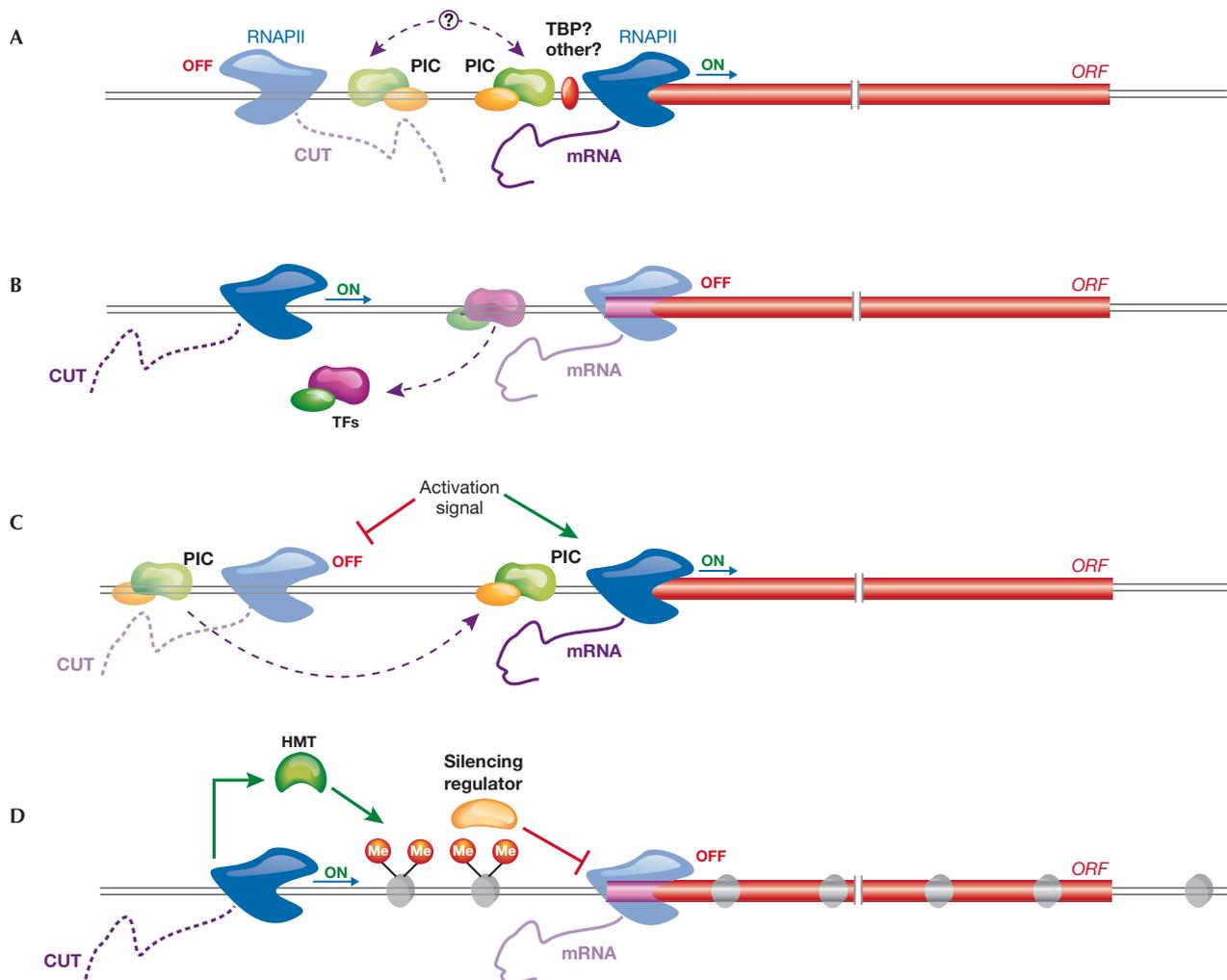


Fig 2 | Possible mechanisms for the regulation of genome expression by non-coding transcription. (A) Bidirectional PARs and mRNAs might originate from different pre-initiation complexes (PICs) and compete for the same pool of transcription factors to initiate transcription. Binding of TBP or other factors might be responsible for directing the balance towards mRNA synthesis. (B) The transcriptional interference mechanism, in which transcription factors (TFs) are displaced from the mRNA promoter by the upstream cryptic transcription, is shown. The *SRG1* cryptic non-coding RNA (ncRNA) interferes with the promoter of the downstream *SER3* gene through this mechanism. (C) Model for start site selection. The CUT and the mRNA have the same promoter but originate from different transcription start sites and compete for the same pool of PIC factors. An example of this type of regulation occurs at the *IMD2* locus. (D) Transcription-induced chromatin modifications, in which cryptic transcription modifies promoter proximal chromatin to attenuate gene expression. The *GAL10–GAL1* locus is regulated through this mechanism; cryptic transcription that originates upstream from the *GAL10–GAL1* promoter induces the methylation of H3K4 and/or H3K36 by the HMTs Set1 and Set2, respectively, and tethers the Rpd3S histone deacetylase complex to attenuate gene expression of the GAL locus. CUT, cryptic unstable transcript; H3, histone H3; HMT, histone methyl transferase; *IMD2*, inosine monophosphate dehydrogenase 2; K, lysine; PAR, promoter-associated non-coding RNA; Rpd3S, reduced potassium dependency 3 small; *SER3*, serine requiring 3; Set1/2, SET-domain-containing 1/2; *SRG1*, *SER3* regulatory gene; TBP, TATA binding protein.

transcript pairs show a strong tendency for co-regulation, confirming that they have the same promoter (Neil *et al*, 2009; Xu *et al*, 2009). Although some of the non-coding transcription events might be inconsequential, it is becoming more and more evident that others are functional, either through the act of transcribing *per se* or, alternatively, through the production of a regulatory transcript. Some examples of functional ncRNAs that have been well characterized are detailed in Table 2.

Transcriptional interference

In some cases, the ncRNA sequence is not conserved, but the promoters and transcription start site are (Carninci *et al*, 2005; Kapranov *et al*, 2007b), which is consistent with the hypothesis that transcription has a function *per se*. In this regard, studies in yeast provided the first example of non-coding transcriptional interference. The transcription of the ncRNA *SRG1* was shown to interfere with the promoter of the downstream *SER3* stress-responsive gene by

Table 2 | Well-characterized eukaryotic long regulatory non-coding RNAs

ncRNA name (organism)	Origin	Function	Regulation	Mechanism	References
Non-coding <i>PHO84</i> (<i>S. cerevisiae</i>)	<i>PHO84</i> , antisense	Regulates <i>PHO84</i>	Ageing/ Rrp6 Rrp6/ Set1	<i>Cis</i> , through HDAC1/2 chromatin modifications <i>Trans</i>	Camblong <i>et al</i> , 2007 Camblong <i>et al</i> , 2009
<i>RTL</i> (<i>S. cerevisiae</i>)	<i>TY1</i> , antisense	Regulates <i>TY1</i> retrotransposons	Xrn1	<i>Trans</i> , through Set1 HMT activity	Berretta <i>et al</i> , 2008
<i>IGS1-R</i> (<i>S. cerevisiae</i>)	rDNA repeats	rDNA stability rDNA copy number control?	Trf4/Rrp6	<i>Cis</i> ? Chromatin structure alteration? Promotion of DNA repair?	Houseley <i>et al</i> , 2007
Non-coding <i>KCSI</i> (<i>S. cerevisiae</i>)	<i>KCSI</i> , antisense	Modulation of Pi signalling	Pi/Pho4	<i>Trans</i> Acts on <i>KCSI</i> translation through RNA–RNA hybrids	Nishizawa <i>et al</i> , 2008
<i>TERRA</i> (<i>S. cerevisiae</i>)	Telomeric repeats	Telomere length control	Rat1	Telomerase inhibition RNA–DNA hybrids?	Luke <i>et al</i> , 2008
<i>TERRA</i> (human, mouse)	Telomeric repeats	Telomeric heterochromatin assembly, telomere length	Effectors for nonsense-mediated decay	RNA–RNA interaction with RNA component of telomerase?	Azzalin <i>et al</i> , 2007; Schoeftner & Blasco, 2008
Non-coding <i>CCND1</i> (human)	<i>CCND1</i>	Inhibition of HAT p300 and CBP, and repression of <i>CCND1</i>	DNA damage signals	<i>Cis</i> , through recruitment of TLS	Wang <i>et al</i> , 2008
<i>DHFR</i> minor transcript (human)	<i>DHFR</i>	Pre-initiation complex dissociation from <i>DHFR</i>	Cell cycle, expressed in quiescent cells	<i>Trans</i> Inhibition of TFIIB Triplex structures with the promoter?	Martianov <i>et al</i> , 2007
<i>Xist</i> (mouse)	Centre of the inactive X chromosome	X-chromosome inactivation	Tsix ncRNA, CTCF chromatin insulator and transcription factor	<i>Cis</i> X-chromosome coating; chromatin modifications; repressive compartment	Reviewed in Chow & Heard, 2009; Payer & Lee, 2008
<i>Air</i> (mouse)	Intron 2 of <i>Igf2r</i>	Silencing of clusters of imprinted genes	Imprinting, DNA methylation	<i>Cis</i> DNA coating? Repressive compartment? Recruitment of G9a HMT? Repressive histone marks?	Nagano <i>et al</i> , 2008
<i>Kcnq1ot1</i> (mouse)	Intron 10 of <i>Kcnq1</i>	Silencing of cluster of imprinted genes	Promoter methylation	<i>Cis</i> DNA coating? DNA methylation, recruitment of PRC2 and G9a HMT? Repressive histone marks?	Pandey <i>et al</i> , 2008
<i>HOTAIR</i> (human)	<i>HOXC</i> locus	Silencing of the <i>HOXD</i> locus	?	<i>Trans</i> Recruitment of PRC2? Repressive histone marks	Rinn <i>et al</i> , 2007

This list of non-coding RNAs (ncRNAs) is not exhaustive and only includes the eukaryotic long ncRNAs for which functional and mechanistic information is available. *Air*, antisense *Igf2r* RNA; CBP, CREB binding protein; *CCND1*, cyclin D1; CTCF, CCCTC-binding factor; *DHFR*, dehydrofolate reductase; HAT, histone acetyl transferase; HMT, histone methyl transferase; *HOTAIR*, HOX antisense intergenic RNA; *IGS1-R*, intergenic sequence 1-R; *Kcnq1ot1*, KCNQ1 overlapping transcript 1; *KCSI*, pKC1 suppressor; *PHO4/84*, phosphate metabolism 4/84; Pi, phosphate inorganic; PRC2, polycomb repressive complex 2; Rat1, ribonucleic acid trafficking 1; rDNA, ribosomal DNA; Rrp6, ribosomal RNA processing 6; *RTL*, antisense LTR; Set1, SET-domain-containing 1; *TERRA*, telomeric-repeat-containing RNAs; TFIIB, transcription factor II B; TLS, translocated in liposarcoma; Trf4, topoisomerase I requiring function 4; Tsix, antisense of *Xist*; *TY1*, transposon in yeast 1; *Xist*, X (inactive)-specific transcript; Xrn1, 5'–3' exoribonuclease 1.

blocking the binding of transcription factors (Martens *et al*, 2004; Fig 2B). In *Schizosaccharomyces pombe*, the transcription of a cascade of ncRNAs through the promoter of the *fbp* gene leads to the opening of the chromatin structure, thereby increasing the accessibility of transcription factors and RNAPII (Hirota *et al*, 2008). A similar event leads to polycomb repressive element (PRE) activation and subsequent chromatin repression in higher eukaryotes (Schmitt *et al*, 2005). The *PHO5* stress gene encodes

promoter-driven sense-oriented CUTs, but also an antisense CUT that originates from the gene 3' end and has been shown to stimulate *PHO5* transcription by increasing the accessibility of the RNAPII to the promoter, probably by contributing to efficient histone eviction (Uhler *et al*, 2007). However, it must be noted that the stimulation of gene expression through antisense transcription does not seem to be a general rule. Indeed, sense–antisense transcript pairs are generally anti-correlated (Neil *et al*,

2009; Xu *et al*, 2009), presumably owing to transcriptional interference or to the generation of inhibitory histone modifications on the coding region.

A different type of regulation involves the transcription of a CUT and an mRNA in tandem and has recently been described at the *S. cerevisiae* *IMD2* and *URA2* loci. In both cases, the transcription of the CUT negatively regulates the expression of the downstream gene. Both the CUT and the mRNA have the same promoter but originate from different transcription start sites and have been proposed to compete for the pre-initiation complex (PIC; Kuehner & Brow, 2008; Fig 2C). In a different—but compatible—model, the transcription of the CUT allows the maintenance of a pool of RNAPII at the gene promoter (Thiebaut *et al*, 2008). Therefore, the CUT transcription would function as an attenuator of gene expression under repressive conditions, and would allow rapid switches to the productive state when the cell receives activation signals. Transcriptional interference and start site selection-dependent regulation are probably quite common given that a number of yeast promoters transcribe a ncRNA immediately upstream from the protein-coding RNA and in the same orientation.

Another mechanism through which cryptic transcription affects gene expression is by influencing the epigenetic state of chromatin. We and others have recently shown the impact that non-coding transcription has on repressive histone modifications at the *GAL10–GAL1* locus in *S. cerevisiae* (Houseley *et al*, 2008; Pinskaya *et al*, 2009). Notably, cryptic transcription—which originates from the coding region of *GAL10* under repressive conditions for *GAL10–GAL1*—has been reported to be responsible for the trimethylation of histone 3 at Lys 36 (H3K36me3) by Set2 (Houseley *et al*, 2008) and H3K4me2/me3 by Set1 (Pinskaya *et al*, 2009) across the entire locus. This leads to the tethering of Rpd3S HDAC and to histone deacetylation (Fig 2D). Pinskaya and co-workers proposed that cryptic transcription inhibits PIC formation and attenuates *GAL1* induction. Although H3K36me3 and H3K4me2/me3 had been reported previously to be associated with actively transcribed chromatin, these recent results highlight their function in repressing the expression of genes that need to be switched on or off rapidly on metabolic change. Similar observations have been reported for other glucose-repressed genes such as *MPH2*, *ICL2* (Houseley *et al*, 2008) or *SUC2* (Pinskaya *et al*, 2009), emphasizing the significance of this type of regulation in the response to changes in nutrient conditions.

Whether the widespread divergent transcription that occurs at eukaryotic promoters has a function remains unclear. It has been hypothesized that the role of bidirectional transcription is to maintain an open chromatin structure at promoters (Xu *et al*, 2009), thereby providing an access platform for transcription factors. Another possibility that has been suggested is that divergent transcription provides a rapidly available pool of RNAPII molecules for the expression of protein-coding mRNAs (Preker *et al*, 2008). Several chromatin modifications have been shown to be associated with the presence of RNAPII. Therefore, divergent transcription could act as an anchor for chromatin remodelers or histone modification complexes. In this regard, high levels of human PROMPTs are associated with promoters that contain many CpG repeats, suggesting a potential role for ncRNA transcription in DNA methylation (Preker *et al*, 2008) and hence in transcriptional repression.

RNA interference: *cis*-effects

Regulatory ncRNAs can act *in cis*, at the site of transcription, or *in trans*, by regulating gene expression at both the transcription site and other genomic loci through various mechanisms (Fig 3). X-chromosome dosage compensation in mammals has been studied extensively and is known to be initiated by a long ncRNA known as *Xist* (Chow & Heard, 2009; Payer & Lee, 2008). *Xist* operates *in cis* by coating the X chromosome and recruiting polycomb-group proteins that induce H3K27me3 and subsequent transcriptional silencing, and has been the only known example of a long regulatory RNA for more than 15 years. Similarly, the long *Air* and *Kcnq1ot1* ncRNAs were recently shown to be necessary for epigenetic gene silencing of paternally imprinted genes in the *Igf2r/Air* and *Kcnq1ot1* clusters, respectively (Nagano *et al*, 2008; Pandey *et al*, 2008). In addition, the human *CCND1* ncRNA has been recently proposed to recruit the *TLS* RNA-binding protein to the transcriptional start site of the *CCND1* gene and to induce allosteric modifications in *TLS* that lead to the inactivation of the CBP and p300 histone acetyl transferases (HATs), thereby indirectly regulating the transcription unit (Wang *et al*, 2008).

Yeast CUTs also have *cis*-acting regulatory functions. In ageing cells, for example, an antisense RNA controls the expression of the *PHO84* metabolic gene (Camblong *et al*, 2007). In those conditions, Rrp6 is delocalized and the stabilization of the CUT—rather than its transcription—is responsible for gene silencing through histone deacetylation. Notably, *cis*-acting RNAs are also involved in the regulation of repetitive elements. A heterochromatin-derived CUT has been suggested to participate in silencing and copy-number control of the rDNA loci in *S. cerevisiae*, thereby implicating cryptic transcripts in genome integrity (Houseley *et al*, 2007; Vasiljeva *et al*, 2008).

Furthermore, the recently identified telomeric-repeat-containing RNAs (TERRA) are also involved in telomere maintenance and genomic stability. In humans, an increase of TERRA levels bound to telomeres induces the loss of entire telomere tracts and genomic instability (Azzalin *et al*, 2007), and *in vitro* experiments suggest that the telomeric RNAs inhibit telomerase activity (Schoeffner & Blasco, 2008). Yeast TERRA are unstable and rapidly degraded by Rat1, but their stabilization leads to defects in telomerase-dependent telomere elongation. It has been proposed that TERRA induce DNA–RNA hybrid formation at telomeres that are responsible for telomerase inhibition and telomere shortening (Luke *et al*, 2008).

RNA interference: *trans*-effects

An essential feature of RNA is its capacity to move inside the nucleus—which allows it to affect distant loci—and between the different cellular compartments, allowing for long-range regulatory effects. Accordingly, ncRNAs can act *in trans* at the transcriptional or post-transcriptional levels. Small double-stranded RNAs have been long known to be involved directly in the regulation of gene expression and/or heterochromatin formation (Carthew & Sontheimer, 2009). For example, in higher eukaryotes, micro RNAs (miRNAs) lead to post-transcriptional gene silencing by stimulating the degradation of target mRNAs or by inhibiting their translation. In fission yeast, siRNAs resulting from the pericentromeric region have been implicated in the nucleation of heterochromatin at centromeres. However, the involvement of other types of non-coding transcript in the modulation of gene expression *in trans* has been recently documented. Antisense ncRNAs can control gene expression

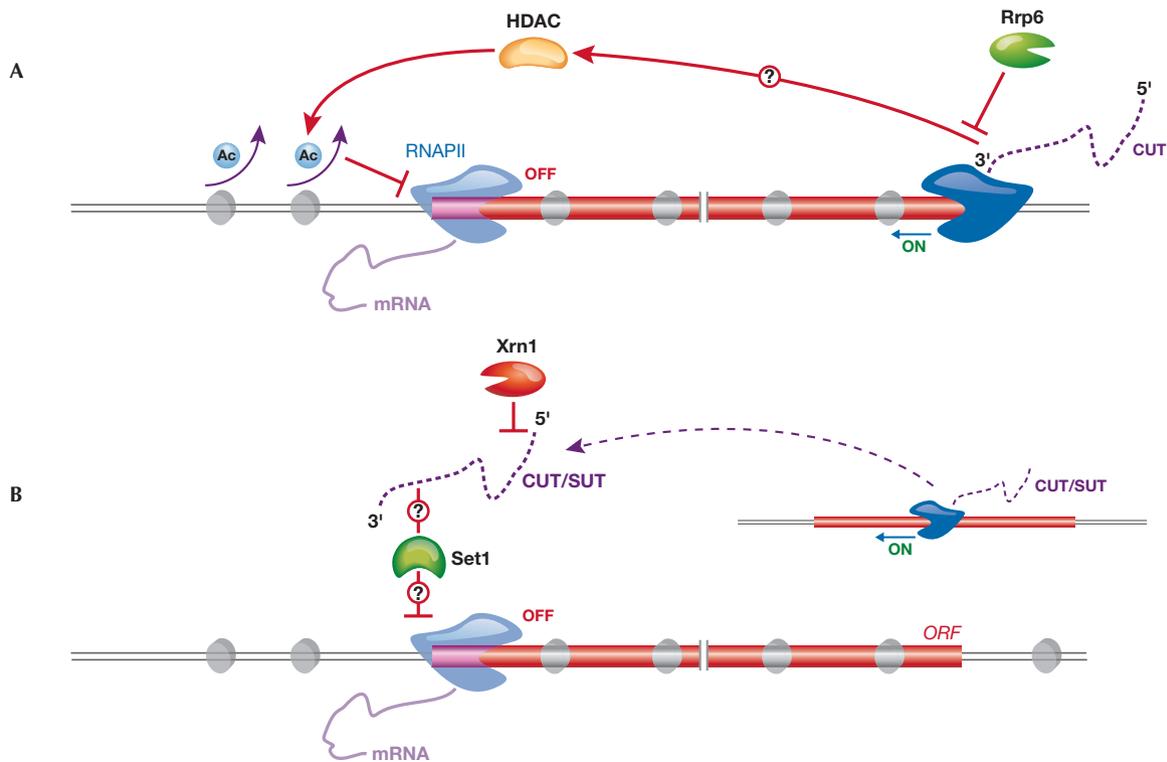


Fig 3 | Models for *cis*- or *trans*-mediated RNA-dependent regulation of gene expression. (A) Regulation *in cis*: when Rnp6 is delocalized or absent, the antisense CUT is stabilized and recruits HDACs, which are responsible for promoter regulation and silencing. This occurs, for example, at the *PHO84* locus. (B) Regulation *in trans*: the CUT, which is transcribed from a distant locus and stabilized, induces the recruitment of the HMT Set1, thereby inhibiting gene transcription. The *RTL* non-coding RNA regulates the *TY1* locus in this manner. CUT, cryptic unstable transcript; HDAC, histone deacetylase complex; HMT, histone methyl transferase; *PHO84*, phosphate metabolism 84; Rnp6, ribosomal RNA processing 6; *RTL*, antisense of LTR; Set1, SET-domain-containing 1; *TY1*, transposon in yeast 1.

post-transcriptionally by inhibiting the translation of protein-coding RNAs, as has been shown at the *KCS1* locus in yeast (Nishizawa *et al*, 2008). In humans, a ncRNA transcribed from a transcription start site that is upstream from the *DHFR* promoter inhibits *DHFR* expression by binding to TFIIB, thereby destabilizing the PIC and hindering transcription (Martianov *et al*, 2007). Another example is the silencing of the *HOX* loci during the development of the *Drosophila* embryo. A *trans*-acting intergenic transcript—known as *HOTAIR*—probably directs the PRC2 methylation complex to a target locus, thereby inducing H3K27me and silencing (Rinn *et al*, 2007). In addition, we have shown that an antisense CUT that starts from an exonic promoter of the yeast *Ty1* retrotransposon controls the mobile elements at the level of transcription and that it can act *in trans* (Berretta *et al*, 2008). The mechanism by which the *TY1* loci are repressed involves Set1 and the associated H3K4me2/me3. We propose that this chromatin mark is necessary for the formation of chromatin domain boundaries between the silenced *TY1* locus and the flanking regions (Berretta *et al*, 2008) or, alternatively, that Set1 marks silenced chromatin as shown for *GAL10–GAL1* (Pinskaya *et al*, 2009). The *Ty1* ncRNA is stabilized in *xrn1* deletion mutants, but regulates the transcription of the retrotransposon, consistent with its *trans*-effect. Interestingly, the *trans*-silencing activity seems not to be restricted to the *Ty1* retro-elements, as the *PHO84* antisense CUT involved in *cis*-silencing of the gene (Camblong *et al*,

2007) has also recently been attributed an additional function in the transcriptional *trans*-silencing of *PHO84*, independently of histone deacetylation (Camblong *et al*, 2009).

Although a number of ncRNAs might act through their secondary structure by promoting conformational changes in regulatory proteins and processes, primary sequence complementarity seems to have a crucial role in RNA-dependent *trans*-regulation. Indeed, the activity of most *trans*-acting ncRNAs is related to the presence of specific sequences, and most of the *trans*-acting RNAs discovered so far are derived from antisense transcription, allowing potential pairing with the mRNA. For example, the *TY1* CUT covers the *TY1* promoter sequence, and the complete sequence of the ncRNA is required for retrotransposition control (Berretta *et al*, 2008). In addition, potent *PHO84* *trans*-silencing is achieved only when long antisense RNAs, which contain sequences complementary to both *PHO84* UAS and 3' end, are produced (Camblong *et al*, 2009). *In vitro* experiments have shown that the *DHFR* ncRNA is able to form RNA–DNA hybrids with the promoter region of *DHFR* (Martianov *et al*, 2007). Also, the *KCS1* ncRNA has been suggested to inhibit translation through RNA–RNA interactions with the mRNA (Nishizawa *et al*, 2008). It is therefore plausible that *trans*-acting RNAs, especially those that have an antisense orientation, could function through sequence complementarity, as described for miRNAs and siRNAs.

Sidebar A | In need of answers

- (i) What is the function of widespread divergent transcription?
- (ii) How is bidirectional transcription regulated?
- (iii) Are mammalian PROMPTs functional?
- (iv) How do *cis*-acting RNAs recruit chromatin-modification factors?
- (v) How do *trans*-acting RNAs induce chromatin modifications?

Concluding remarks

This overview of non-coding, widespread transcription in eukaryotes highlights the complex pattern of transcription-mediated regulation of genome plasticity. RNA molecules have several features that make them ideal regulatory molecules; they are mobile and are rapidly synthesized and degraded, allowing a fast turnover and adaptation to environmental conditions. RNA binds to regulatory or structural proteins, but also to DNA and RNA through sequence complementarity. It can therefore target proteins to chromatin or to other RNA molecules and induce protein conformational changes. The RNA processing machinery is also involved in RNA-mediated control. One possibility is that Rrp6 mainly controls *cis*-acting transcripts, whereas Xrn1 would have a preferential role on *trans*-acting RNAs, which is consistent with the subcellular localization of these proteins. Further insight into the regulation of the RNA decay pathway will help to understand the physiological role of cryptic transcription. Indeed, Rrp6 has been shown to be removed from chromatin in ageing cells (Camblong *et al*, 2007) and high lithium concentrations downregulate Xrn1 (Dichtl *et al*, 1997), leading to the accumulation of the *Ty1* ncRNA (J. Berretta, unpublished results).

The process of transcription *per se* also has a very important role in the modulation of gene expression, mainly by changing the state of chromatin. However, the precise role of nascent transcripts in targeting or regulating chromatin factors remains to be elucidated. Transcription-related regulation mainly controls the expression of genes involved in response to environmental conditions, and is probably necessary for rapid switches between repressive and active chromatin states, as well as for the fine-tuning of gene expression.

Many questions remain unanswered about the regulation, functions and mechanisms of non-coding RNAs (Sidebar A). Spurious RNAs and cryptic transcription have been shown to participate in different pathways, ranging from gene expression modulation to telomere silencing control, and further investigation will be necessary to unveil how these molecules mediate their effects.

ACKNOWLEDGEMENTS

We are grateful to A. Jacquier and L. Steinmetz for pertinent discussion and critical reading of the manuscript. We also thank M. Kwapisz, E. Van Dijk and M. Pinskaya for useful comments and suggestions, and F. Stutz for personal communications and discussions on unpublished data. Work in the laboratory of A.M. is supported by Human Frontier Science Program Organization, Association pour la Recherche sur le Cancer, Fondation pour la Recherche Médicale and Agence Nationale pour la Recherche (REGULncRNA). A.M. is a European Molecular Biology Organization Young Investigator.

REFERENCES

Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**: 744–747

Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ (2008) A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice. *Mol Cell* **31**: 785–799

Arigo JT, Eyer DE, Carroll KL, Corden JL (2006) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell* **23**: 841–851

Azzalin CM, Reichenbach P, Khoraiuli L, Giulotto E, Lingner J (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**: 798–801

Berretta J, Pinskaya M, Morillon A (2008) A cryptic unstable transcript mediates transcriptional *trans*-silencing of the *Ty1* retrotransposon in *S. cerevisiae*. *Genes Dev* **22**: 615–626

Bertone P *et al* (2004) Global identification of human transcribed sequences with genome tiling arrays. *Science* **306**: 2242–2246

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**: 1089–1103

Camblong J, Iglesias N, Fickentscher C, Dieppl G, Stutz F (2007) Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* **131**: 706–717

Camblong J, Beyrouthy N, Guffanti E, Schlaepfer G, Steinmetz LM, Stutz F (2009) *Trans*-acting antisense RNAs mediate transcriptional gene cosuppression in *S. cerevisiae*. *Genes Dev* **23**: 1534–1545

Carninci P *et al* (2005) The transcriptional landscape of the mammalian genome. *Science* **309**: 1559–1563

Carninci P *et al* (2006) Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* **38**: 626–635

Carrozza MJ *et al* (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581–592

Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell* **136**: 642–655

Cheng J *et al* (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **308**: 1149–1154

Chow J, Heard E (2009) X inactivation and the complexities of silencing a sex chromosome. *Curr Opin Cell Biol* **21**: 359–366

Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**: 1845–1848

David L, Huber W, Granovskaia M, Toedling J, Palm CJ, Bofkin L, Jones T, Davis RW, Steinmetz LM (2006) A high-resolution map of transcription in the yeast genome. *Proc Natl Acad Sci USA* **103**: 5320–5325

Dichtl B, Stevens A, Tollervey D (1997) Lithium toxicity in yeast is due to the inhibition of RNA processing enzymes. *EMBO J* **16**: 7184–7195

Dutrow N, Nix DA, Holt D, Milash B, Dalley B, Westbroek E, Parnell TJ, Cairns BR (2008) Dynamic transcriptome of *Schizosaccharomyces pombe* shown by RNA–DNA hybrid mapping. *Nat Genet* **40**: 977–986

Guttman M *et al* (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**: 223–227

He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW (2008) The antisense transcriptomes of human cells. *Science* **322**: 1855–1857

Hirota K, Miyoshi T, Kugou K, Hoffman CS, Shibata T, Ohta K (2008) Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. *Nature* **456**: 130–134

Houseley J, Kotovic K, El Hage A, Tollervey D (2007) Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. *EMBO J* **26**: 4996–5006

Houseley J, Rubbi L, Grunstein M, Tollervey D, Vogelauer M (2008) A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Mol Cell* **32**: 685–695

Kaplan CD, Laprade L, Winston F (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301**: 1096–1099

Kapranov P *et al* (2007a) RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**: 1484–1488

Kapranov P, Willingham AT, Gingeras TR (2007b) Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet* **8**: 413–423

Katayama S *et al* (2005) Antisense transcription in the mammalian transcriptome. *Science* **309**: 1564–1566

Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B (2005) A high-resolution map of active promoters in the human genome. *Nature* **436**: 876–880

- Kuehner JN, Brow DA (2008) Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol Cell* **31**: 201–211
- LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**: 713–724
- Lee A, Hansen KD, Bullard J, Dudoit S, Sherlock G (2008) Novel low abundance and transient RNAs in yeast revealed by tiling microarrays and ultra high-throughput sequencing are not conserved across closely related yeast species. *PLoS Genet* **4**, e1000299.
- Li C, Mueller JE, Bryk M (2006) Sir2 represses endogenous polymerase II transcription units in the ribosomal DNA nontranscribed spacer. *Mol Biol Cell* **17**: 3848–3859
- Li L, Wang X, Stolc V, Li X, Zhang D, Su N, Tongprasit W, Li S, Cheng Z, Wang J, Deng XW (2006) Genome-wide transcription analyses in rice using tiling microarrays. *Nat Genet* **38**: 124–129
- Luke B, Panza A, Redon S, Iglesias N, Li Z, Lingner J (2008) The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in *Saccharomyces cerevisiae*. *Mol Cell* **32**: 465–477
- Martens JA, Laprade L, Winston F (2004) Intergenic transcription is required to repress the *Saccharomyces cerevisiae* *SER3* gene. *Nature* **429**: 571–574
- Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A (2007) Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **445**: 666–670
- Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF (2008) A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res* **18**: 1073–1083
- Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**: 1344–1349
- Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**: 1717–1720
- Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**: 1038–1042
- Nishizawa M, Komai T, Katou Y, Shirahige K, Ito T, Toh EA (2008) Nutrient-regulated antisense and intragenic RNAs modulate a signal transduction pathway in yeast. *PLoS Biol* **6**: 2817–2830
- Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* **32**: 232–246
- Payer B, Lee JT (2008) X chromosome dosage compensation: how mammals keep the balance. *Annu Rev Genet* **42**: 733–772
- Pinskaya M, Gourvenec S, Morillon A (2009) H3 lysine 4 di- and trimethylation deposited by cryptic transcription attenuates promoter activation. *EMBO J* **28**: 1697–1707
- Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, Schierup MH, Jensen TH (2008) RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**: 1851–1854
- Rinn JL *et al* (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**: 1311–1323
- Schmitt S, Prestel M, Paro R (2005) Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev* **19**: 697–708
- Schoefer S, Blasco MA (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* **10**: 228–236
- Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA (2008) Divergent transcription from active promoters. *Science* **322**: 1849–1851
- Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA (2006) Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell* **24**: 735–746
- Stolc V *et al* (2004) A gene expression map for the euchromatic genome of *Drosophila melanogaster*. *Science* **306**: 655–660
- Taft RJ *et al* (2009) Tiny RNAs associated with transcription start sites in animals. *Nat Genet* **41**: 572–578
- Thiebaud M, Kisseleva-Romanova E, RougE-maille, M, Boulay J, Libri D (2006) Futile cycle of transcription and nuclear degradation of cryptic unstable transcripts: a role for the nrd1–nab3 pathway in genome surveillance. *Mol Cell* **23**: 853–864
- Thiebaud M, Colin J, Neil H, Jacquier A, Seraphin B, Lacroute F, Libri D (2008) Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in *S. cerevisiae*. *Mol Cell* **31**: 671–682
- Thompson DM, Parker R (2007) Cytoplasmic decay of intergenic transcripts in *Saccharomyces cerevisiae*. *Mol Cell Biol* **27**: 92–101
- Uhler JP, Hertel C, Svestrup JQ (2007) A role for noncoding transcription in activation of the yeast *PHO5* gene. *Proc Natl Acad Sci USA* **104**: 8011–8016
- Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol* **3**: e189
- Vasiljeva L, Kim M, Terzi N, Soares LM, Buratowski S (2008) Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. *Mol Cell* **29**: 313–323
- Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE Jr, Hieter P, Vogelstein B, Kinzler KW (1997) Characterization of the yeast transcriptome. *Cell* **88**: 243–251
- Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, Tempst P, Rosenfeld MG, Glass CK, Kurokawa R (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* **454**: 126–130
- Whitehouse I, Rando OJ, Delrow J, Tsukiyama T (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**: 1031–1035
- Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bahler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**: 1239–1243
- Wyers F *et al* (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**: 725–737
- Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Munster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**: 1033–1037



Antonin Morillon & Julia Berretta