OPINION

Regulatory feedback from nascent RNA to chromatin and transcription

Lenka Skalska, Manuel Beltran-Nebot, Jernej Ule and Richard G. Jenner

Abstract | Transcription and chromatin function are regulated by proteins that bind to DNA, nucleosomes or RNA polymerase II, with specific non-coding RNAs (ncRNAs) functioning to modulate their recruitment or activity. Unlike ncRNAs, nascent pre-mRNA was considered to be primarily a passive player in these processes. In this Opinion article, we describe recently identified interactions between nascent pre-mRNAs and regulatory proteins, highlight commonalities between the functions of nascent pre-mRNA and nascent ncRNA, and propose that both types of RNA have an active role in transcription and chromatin regulation.

Regulators of transcription and chromatin function are localized to genes by binding to DNA, histones or RNA polymerase II (Pol II). That RNA can also have a role in the recruitment of transcription regulators in mammalian cells was first demonstrated for the transactivation response element (TAR), which is an RNA stem-loop formed at the 5′ end of nascent HIV transcripts that recruits the viral protein transactivator of transcription (Tat) and the cellular positive transcription elongation factor b (P-TEFb)[1]. Studies of long non-coding RNAs (lncRNAs), including X-inactive-specific transcript (*XIST*), *KCNQ1* opposite strand/antisense transcript 1 (*KCNQ1OT1*) and *HOX* transcript antisense RNA (*HOTAIR*), which associate with Polycomb repressive complex 2 (PRC2)[2–5], revealed that such specialized cellular RNAs could directly bind to regulatory proteins and modulate their recruitment to genes. There is also evidence that other types of ncRNAs, such as those transcribed from enhancers, also contribute to gene regulation, in either their nascent or mature forms[6–9].

Recent studies have revealed that the interaction between RNA and regulators of chromatin and of transcription is not limited to specialized ncRNA species. Instead, several ncRNA-binding chromatin regulators interact extensively with pre-mRNA[10,11]. For instance, PRC2 directly interacts with nascent RNAs at essentially all active genes, without any preference for lncRNAs or pre-mRNAs[10]. Thus, some of the seemingly ncRNA-specific properties are emerging as more general properties of nascent RNA transcripts, including pre-mRNAs. ncRNA-specific properties are discussed in Supplementary information S1 (box).

Here, we present evidence from recent studies suggesting that pre-mRNA has an active role in regulating transcription and chromatin function. Focusing on mammals and on non-RNAi-based mechanisms, we first discuss the role of specific RNA elements in coupling RNA processing with transcription elongation and chromatin modification. We then detail how more promiscuous interactions between nascent RNA and transcription factors and chromatin-modifying complexes can promote or repress their function on chromatin. We conclude by discussing functional commonalities between different types of nascent RNAs and considering the relationship between nascent RNA and higher-order chromatin structure.

Roles of specific RNA elements

Nascent pre-mRNAs can contain specific sequences and structures that regulate Pol II pausing and chromatin modification. Transcription elongation factors bind to sequences at the 5′ end of cellular pre-mRNAs (FIG. 1a), and splice sites influence the Pol II elongation rate and chromatin modification across the gene body (FIG. 1b). At the 3′ end of genes, Pol II pausing occurs after recognition of the polyadenylation site (PAS) by cleavage and polyadenylation factors[12] and owing to the formation of RNA–DNA hybrids known as R-loops (BOX 1; FIG. 1c; discussed in more detail in REFS 12,13).

Nascent RNA at the 5′ end of genes. The recruitment of P-TEFb from the inhibitory 7SK ribonucleoprotein (RNP) complex to the HIV TAR RNA[1] serves as a paradigm for the role of nascent RNA in regulating transcription elongation at the 5′ end of genes. A related mechanism seems to be in operation at cellular genes. Exonic splicing-enhancer sequences at the 5′ end of pre-mRNAs can also recruit P-TEFb, together with Ser/Arg-rich splicing factor 2 (SRSF2)[14] (FIG. 1a). Similarly, the negative elongation factor (NELF) complex, which is a regulator of Pol II promoter-proximal pausing, directly binds — through its NELFE subunit — to both HIV TAR[15] and to sequence elements at the 5′ end of nascent cellular RNAs[16,17]. The SPT5 subunit of the DRB sensitivity-inducing factor (DSIF) complex was also shown *in vitro* to contact the nascent RNA as it emerges from Pol II[17], and it has recently been found to interact specifically with the 5′ end of nascent pre-mRNAs at all active genes (T. Henriques, B. S. Scruggs, R. A. Flynn, M. O. Inouye, G. W. Muse, A. Burkholder, C. A. Lavender, D. C. Fargo, H. Y. Chang and K. Adelman, personal communication). Thus, the recruitment of transcription elongation factors to RNA at the 5′ end of genes seems to be a broadly acting cellular mechanism for regulating promoter-proximal pause release.

Coupling of splicing and transcription elongation. Spliceosome assembly, and often also splicing, occurs co-transcriptionally at most genes, which allows crosstalk between the nascent mRNA and Pol II elongation and chromatin modification (reviewed in REFS 18–20). Pol II elongation rate, Pol II carboxy-terminal domain (CTD)

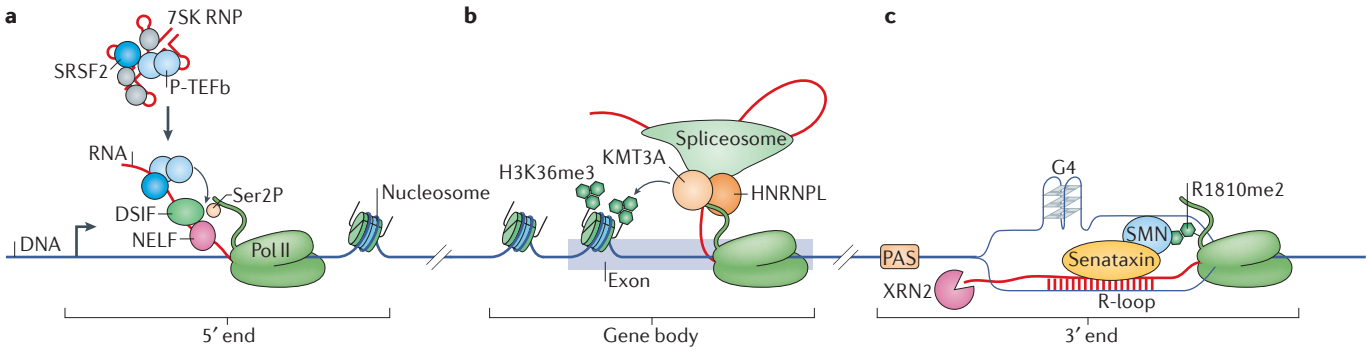

Figure 1 | **Nascent RNA couples RNA processing with transcription elongation and chromatin modification at specific positions along the gene. a** | Positive transcription elongation factor b (P-TEFb) can be recruited to the 5′ end of genes from the inhibitory 7SK ribonucleoprotein (RNP) through the interaction of the splicing factor Ser/Arg-rich splicing factor 2 (SRSF2) with the pre-mRNA. P-TEFb can then phosphorylate the RNA polymerase II (Pol II) carboxy-terminal domain (CTD) at Ser2 (Ser2P). The complexes negative elongation factor (NELF) and DRB sensitivity-inducing factor (DSIF) also directly bind to the nascent mRNA at the 5′ end of genes, through the NELFE and SPT5 subunits (not shown), respectively. **b** | The methyltransferase KMT3A is recruited to specific locations within the gene through mechanisms dependent on splicing and interactions with heterogeneous nuclear RNP L (HNRNPL), contributing to the enrichment of histone H3 Lys36 trimethylation (H3K36me3) at exons. **c** | RNA–DNA hybrids (termed R-loops) induce Pol II pausing and transcription termination at the 3′ end of genes. The nascent RNA can hybridize with the unwound template DNA and displace the non-template strand, which can be stabilized by forming G-quadruplex structures (G4). The helicase senataxin is recruited to R-loops by the binding of survival motor neuron protein (SMN) to the Pol II CTD dimethylated at Arg1810 (R1810me2)[80] or through BRCA1 (REF. 81) (not shown). Senataxin resolves these structures, which promotes 5′–3′ exoribonuclease 2 (XRN2)-mediated Pol II transcription termination. PAS, polyadenylation site.

modifications, chromatin modifications and nucleosome density all influence alternative splicing. Reciprocally, splice-site sequences in the nascent pre-mRNA recruit the spliceosome, and Pol II pauses at these sequences, thereby enhancing splicing fidelity. This pausing was first reported in budding yeast, in which the rate of Pol II elongation was found to generally decrease at 3′ splice sites[21]. The pausing was abrogated by both pharmacological inhibition of splicing factors and by mutation of the splice site or branch point sequence[21,22]. The effect of the branch point mutation was rescued by complementary mutations in the U2 small nuclear RNA, which indicates that Pol II pausing depends on pre-spliceosome formation[21]. Genome-wide profiling indicates that a similar mechanism might operate in higher eukaryotes[23–27]. Measurement of Pol II elongation rate in *Drosophila melanogaster* using precision nuclear run-on sequencing (PRO-seq)[27] and in humans using global run-on sequencing (GRO-seq)[24] revealed a reduced rate at exons, with exon density being one of the strongest predictors of elongation rate across the whole gene[24,25,27]. This apparent stalling is dependent on splicing, as pausing occurs only at retained exons and not at skipped exons[27]. Profiling of Pol II kinetics at nucleotide resolution in human cells using native elongating transcript sequencing (NET-seq) suggests that Pol II pauses at both 5′ and 3′ splice junctions[23,26]; however, it remains possible that part of this signal reflects the presence of co-precipitating splicing intermediates. Therefore, further studies will be necessary to confirm the role of the splicing machinery in regulating Pol II processivity.

Coupling of splicing and chromatin modification. Coupling of RNA processing with chromatin modification was initially suggested by the coincidence of certain changes in chromatin with splice sites. Trimethylation of histone H3 at Lys36 (H3K36me3) is associated with exonic sequences and correlates with the level of gene expression and exon inclusion (reviewed in REF. 19). Splice sites are required to establish and maintain H3K36me3 (REFS 28,29), and pharmacological inhibition of U2 small nuclear RNP (snRNP), or direct depletion of its component splicing factor 3B subunit 3 (SF3B3; also known as SAP130), results in loss of H3K36me3 enrichment at exons[28–30]. Furthermore, heterogeneous nuclear RNP L, which is a sequence-specific RNA-binding protein that binds to CA-rich RNA motifs, is associated with the H3K36me3 methyltransferase KMT3A (also known as SETD2)[31]. These data indicate that splice sites and other sequence elements on nascent RNA are important for coupling the recruitment of KMT3A to elongating Pol II[28] (FIG. 1b).

Blocking of transcription repressors

In addition to regulating transcription and chromatin modifications through the presence of specific sequences and structures, nascent RNA serves as a more general signal for gene activity by blocking the function of chromatin modifiers that would otherwise serve to repress transcription.

PRC2. PRC2 associates with CpG islands at genes that regulate development, where it methylates H3K27 to maintain gene repression. The identification of lncRNAs that bind to PRC2 and modulate its association with chromatin[2–5] led to models in which PRC2 recognizes a specific set of lncRNAs, which then direct it to specific sites on chromatin. Various RNAs that co-precipitated with PRC2 in mouse embryonic stem cells were identified by native RNA immunoprecipitation and sequencing (RIP–seq)[32]. Surprisingly, most of these RNAs were found to be protein-coding transcripts[33]. Using photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP), which is a more stringent technique that identifies only direct cellular protein–RNA interactions, a specific set of 774 nascent 'ezRNAs' (named after the Enhancer of zeste homologue 2 (EZH2) subunit of PRC2) were found to be directly bound by PRC2 (REF. 34). These RNAs were proposed to block PRC2 enzymatic

activity but not the association of PRC2 with chromatin, thereby poising the genes for future silencing by PRC2 (REFS 34,35). Expanding on this, a higher sensitivity analysis of RNA binding by PRC2, using individual-nucleotide-resolution crosslinking and immunoprecipitation (iCLIP), demonstrated that PRC2 interacts with nascent, unspliced RNA at essentially all active genes, of which the previously characterized ezRNAs were revealed to be the most highly expressed and longest (hence, experimentally the most identifiable) transcripts and thus just the tip of the PRC2–RNA interactome[10]. This iCLIP analysis also demonstrated that PRC2 has no preference for lncRNAs or any other RNA species, which is consistent with *in vitro* RNA-binding data showing that the complex binds to RNA promiscously[33,36].

What is the function of promiscuous PRC2 binding to nascent RNA? Experiments that degraded the RNA in cells revealed that loss of RNA leads to increased interaction of PRC2 with chromatin at active genes[10]. Reciprocally, release of PRC2 from chromatin increases its interaction with RNA[10]. This suggests that, at each gene, PRC2 binding to chromatin and PRC2 binding to RNA are mutually antagonistic processes: the higher the amount of nascent RNA, the lower the amount of PRC2 that is able to bind to chromatin, which protects active genes from inappropriate silencing by PRC2 (FIG. 2a). Consistent with nascent RNA having an antagonistic effect on chromatin binding by PRC2, inhibition of Pol II induces PRC2 binding to chromatin at active genes[37], forced early termination of Pol II through the insertion of an upstream PAS increases H3K27me3 levels at the promoter[35], and insertion of a promoter and enhancer next to CpG islands blocks PRC2 recruitment[38]. Furthermore, RNA competes with nucleosomes for binding to PRC2 *in vitro*[10], and this prevents H3K27 methylation[35,39,40]. A role for nascent RNA in modulating the binding of PRC2 to chromatin is also indicated by the ability of R-loops to antagonize the association of PRC2 with chromatin[41]. PRC2 has also been reported to interact with RNA-binding protein fox-1 homologue 2 (RBFOX2), the depletion of which reduces PRC2 binding to chromatin at bivalent genes[42]. Taken together, these studies support a model in which nascent RNA competes with chromatin for PRC2 binding, preventing PRC2 recruitment to chromatin at transcriptionally active genes and evicting PRC2 from chromatin on gene activation.

Antagonizing other chromatin regulators. Nascent RNA also acts to block the activity of other modifying complexes on chromatin. Using native RIP and electrophoretic mobility shift assays, DNA (cytosine-5)-methyltransferase 1 (DNMT1) was found to interact with a non-polyadenylated nascent RNA transcribed across the CCAAT/enhancer-binding protein alpha (*CEBPA*) gene[43]. RIP–seq and formaldehyde RIP–seq (fRIP–seq) revealed that DNMT1 associates with nascent RNA at thousands of genes, and that DNMT1 is primarily associated with pre-mRNA rather than with lncRNAs[11]. The degree of DNMT1–RNA association is negatively correlated with promoter DNA methylation *in vivo*[11,43] and, consistent with this, transcription inhibits DNMT1 methyltransferase activity *in vitro*[43]. DNMT3A also binds to nascent RNA, which inhibits its enzymatic activity[44,45], suggesting that inhibition by nascent RNA may be a general feature of DNMTs (FIG. 2a). The RNAs bound by DNMT1 and DNMT3A have been described to be members of a distinct class of non-polyadenylated transcripts termed 'extracoding RNAs' (REF. 45). However, many of these RNAs resemble nascent pre-mRNAs, which are readily detected in the non-polyadenylated fraction[46], and this suggests that DNMTs have a similar RNA-binding profile to that of PRC2.

RNA also has an antagonistic effect on the H3K9 methyltransferase G9a (also known as EHMT2). The lncRNA *ROR* competes with DNA for G9a binding, but this effect is also observed using total RNA, suggesting, as for PRC2, that RNA acts as a general competitive agent for chromatin binding[47]. Using fRIP–seq, several other chromatin repressors have been found to associate with mRNA, including Lys-specific histone demethylase 1 (LSD1; also known as KDM1A; which also binds to the lncRNA *HOTAIR*[2]), chromobox homologue 3 (CBX3) and the nucleosome remodelling and deacetylase (NuRD) complex components histone deacetylase 1 (HDAC1) and chromodomain-helicase-DNA-binding protein 4 (CHD4), all of which preferentially bind to mRNA over lncRNA[11]. Nascent RNA may therefore generally prevent the association of repressive factors with chromatin, thus maintaining a chromatin environment conducive to transcription.

Recruiting activators

In addition to interacting with repressive factors, nascent pre-mRNA interacts with transcription and chromatin regulators that activate gene expression. Such interactions were originally discovered in studies of ncRNAs termed enhancer RNAs (eRNAs), activatory lncRNAs or ncRNA-activating (ncRNA-a), which have been suggested to bring activating factors to the promoters of neighbouring protein-coding genes[48–51]. The activatory lncRNA *HOXA* transcript at the distal tip (*HOTTIP*) binds to WD

Box 1 | R-loops regulate transcription termination and chromatin modification

Nascent mRNA can hybridize with the underwound template DNA upstream of RNA polymerase II (Pol II), thereby displacing the non-template strand (which can be stabilized by forming intramolecular G-quadruplex structures)[12,13] (FIG. 1c). These RNA–DNA hybrid structures, termed R-loops, have recently been demonstrated to regulate Pol II processivity and chromatin modification at a subset of genes. R-loops formed at the 3′ end of genes cause Pol II to pause and subsequently function in transcription termination[71,72] (FIG. 1c). R-loops at the 3′ end of genes can also act to promote chromatin condensation. R-loops can induce antisense transcription at their site of formation, as well as the recruitment of the RNA-induced silencing complex (RISC)[73]. The R-loops and RISC are necessary for the subsequent recruitment of G9a and of heterochromatin protein 1γ[73]. R-loops additionally promote phosphorylation of histone H3 at Ser10, which is also associated with chromatin condensation, although the mechanism underlying this is not yet understood[74].

R-loops formed at the 5′ regions of genes are associated with a subset of CpG island-containing promoters that exhibit a strong GC-skew[72]. Such promoters exhibit elevated levels of transcriptionally active chromatin markers, such as DNase I hypersensitivity and histone H3 Lys4 trimethylation[75–77], and are protected from DNA (cytosine-5)-methyltransferase 3B (DNMT3B)-catalysed DNA methylation. A direct role for R-loops in chromatin modification is suggested by the binding to chromatin of the histone acetyltransferase complex between TIP60 (also known as KAT5) and p400 (also known as EP400) at sites of R-loop formation and by its loss following overexpression of the R-loop-resolving enzyme RNase H1 (REF. 41). By contrast, RNase H1 overexpression was found to increase binding to chromatin of the transcriptionally repressive Polycomb repressive complex 2 (PRC2)[41], consistent with the antagonistic effect of transcription activity on PRC2 chromatin association[10,37].

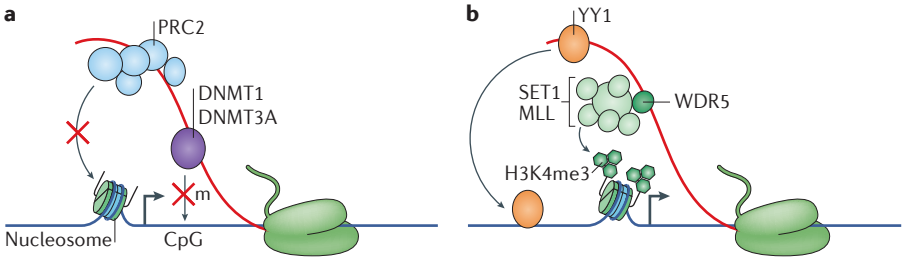

Figure 2 | **Nascent RNA modulates the association of regulatory factors with chromatin to maintain gene activity. a** | Nascent RNA can compete with chromatin for binding of repressive chromatin modifiers, such as Polycomb repressive complex 2 (PRC2), which methylates histone H3 at Lys27, and DNA (cytosine-5)-methyltransferase 1 (DNMT1) and DNMT3A, which primarily methylate the DNA at CpG dinucleotides. **b** | Interaction of the transcription factor yin and yang 1 (YY1) with nascent RNA facilitates its transfer to chromatin. Similarly, the interaction of WD repeat-containing 5 (WDR5), which is a component of the histone Lys methyltransferase complexes SET1 and myeloid/lymphoid or mixed-lineage leukaemia (MLL), with nascent RNA facilitates their transfer to chromatin and trimethylation of histone H3 at Lys4 (H3K4me3), thereby forming a positive-feedback loop that promotes gene expression.

repeat-containing 5 (WDR5), which is a component of the SET1 and myeloid/lymphoid or mixed-lineage leukaemia (MLL) histone Lys4 methyltransferase complexes[48]. Chromatin looping brings the *HOTTIP* gene and RNA close to genes at the 5′ end of the *HOXA* locus, thereby allowing H3K4 trimethylation and gene activation[48]. Unbiased identification of WDR5-associated RNAs using RNA–protein immunoprecipitation in tandem with sequencing (RIPiT–seq)[52] and fRIP–seq[11] revealed that WDR5 is also associated with pre-mRNA and mRNA and that this positively correlates with H3K4 dimethylation (H3K4me2) and H3K4me3 levels in *cis*[11]. The histone acetyltransferase p300/CBP-associated factor (PCAF; also known as KAT2B) was also found by fRIP–seq to associate with RNA in a pattern that is positively correlated with active chromatin modifications[11]. These data suggest that nascent RNA can act as a general binding platform for transcription-activating chromatin-modifying complexes. Why the chromatin recruitment of some complexes, such as MLL, is promoted by nascent RNA, whereas the recruitment of other complexes, such as PRC2, is antagonized by it, is discussed in BOX 2.

Shared functions of nascent pre-mRNAs and eRNAs. The transcription factor yin and yang 1 (YY1) was originally found to bind to the mature lncRNA *XIST*[53]. Using CLIP, YY1 was found to also interact with nascent pre-mRNA and nascent eRNA[54]. In contrast to PRC2, which becomes associated with chromatin on RNA degradation, YY1 is lost from chromatin when permeabilized cells are incubated with RNaseA[54]. Furthermore, tethering of RNA to enhancer DNA increased YY1 binding to chromatin[54]. This suggests that RNA functions to 'trap' YY1 near DNA, thereby increasing the local concentration of YY1 and promoting its subsequent loading onto neighbouring DNA[54] (FIG. 2b). This may create a positive-feedback loop in which transcription factors such as YY1 stimulate the transcription of nascent RNAs, which then retain the transcription factors locally[54]. Thus, eRNAs might function to regulate the binding of regulatory proteins to the enhancer DNA, in addition to having a regulatory role at neighbouring protein-coding genes. This suggests that the binding of transcription regulators to nascent pre-mRNA and to nascent eRNA is to a degree functionally equivalent, with both types of RNAs helping to maintain the activity of the regulatory elements that are adjacent to them, be that a gene promoter or a gene enhancer (FIG. 3a).

Regulatory crosstalk

Regulatory interactions between ncRNA species and protein-coding genes have primarily been considered to be unidirectional, with the ncRNAs regulating the expression of nearby protein-coding genes[48–51]. However, the ability of YY1 to bind to both nascent pre-mRNAs and eRNAs and the promiscuous nature of RNA binding by chromatin regulators such as PRC2 suggest that pre-mRNAs may have effects similar to those of nascent ncRNAs in regulating neighbouring enhancers and genes (FIG. 3a). In support of this, when the promoter of the activity-regulated cytoskeleton-associated (*ARC*) gene is deleted, the production of eRNA from *ARC* enhancers ceases[55]. Furthermore, it was recently found that both lncRNAs and protein-coding genes can regulate the transcription of adjacent genes in *cis*[56]. Although this phenomenon can be the result of gene promoters acting as enhancers for adjacent genes, it can also be dependent on gene transcription or splicing of the nascent transcript[56].

Impact of chromatin looping. The regulatory roles of eRNAs and activatory lncRNAs are thought to be mediated by chromatin looping, which brings them close to their target genes[8,57]. Thus, chromatin looping may allow any form of nascent RNA to contribute to gene regulation in *cis*. By bringing active genes and active regulatory elements in close proximity, chromatin looping would increase the local RNA concentration, and this may potentiate retention of activating proteins such as YY1 and MLL as well as help to outcompete chromatin for binding by repressive factors such as PRC2, thereby keeping the

Box 2 | **Promotion versus repression of regulatory factor activity by nascent RNA**

Why does nascent RNA promote the activity of some factors on chromatin but antagonize others? The answer may lie in the relative affinities of proteins for RNA versus DNA (or chromatin). DNA (cytosine-5)-methyltransferase 1 (DNMT1) has higher affinity for RNA than for DNA[43], suggesting that, if present, RNA may be able to outcompete the DNA for DNMT1 binding. By contrast, yin and yang 1 (YY1) has a higher affinity for DNA than for RNA[54]; thus, its transfer from RNA to DNA will be energetically favourable. If RNA- and DNA-binding (or nucleosome-binding) surfaces overlap, this would lead to mutually exclusive RNA and chromatin binding. Conversely, a protein or protein complex with independent RNA- and DNA-binding domains could be recruited to chromatin by RNA and DNA acting cooperatively. RNA binding can also allosterically modify the ability of proteins to bind chromatin, and enhance or inhibit their catalytic activity[78,79].

The rate of nascent RNA synthesis might also influence its effect on chromatin. RNA polymerase II (Pol II) transcription rates vary between genes and within genes[12,18]. Proteins that bind to the RNA when it is being rapidly polymerized might be quickly taken away from their target sites on chromatin, whereas binding of a protein to RNA that is attached to a stalled Pol II could provide more time for the protein to interact with the chromatin nearby. Indeed, Pol II slows down at YY1-binding sites[26], and this could enhance the capacity of RNA to 'trap' YY1 near chromatin at these sites[54].

chromatin neighbourhood active (FIG. 3b). For some RNA-binding factors, the local concentration of RNA in three-dimensional space, rather than the rate of transcription of nascent RNA at a single locus, could be the best predictor of their degree of chromatin binding.

Formation of chromatin loops. Depletion of nuclear RNA has long been known to disrupt normal chromatin morphology and nuclear organization[58], and long-lived ncRNAs, such as *XIST* and repeats-derived RNAs, have a crucial role in maintaining chromatin architecture[59,60]. Accumulating evidence also indicates that transcription has a role in the formation of higher-order chromatin structures. Depletion of ncRNAs that bind the Mediator complex reduces chromatin looping between genes and enhancers[49]. Consistent with the colocalization of Mediator and cohesin and their shared role in anchoring short-range contacts between promoters and enhancers[61,62], fRIP–seq studies indicate that cohesin also associates with mRNA[11]. Furthermore, CLIP data demonstrate that CCCTC-binding factor (CTCF), which binds to chromatin insulator sites, also directly binds to RNA in a seemingly promiscuous fashion[63,64]. These studies suggest a possible role for nascent RNA in the regulation of chromatin looping.

Nuclear bodies. Nascent pre-mRNAs can also affect nuclear organization by initiating the formation of nuclear bodies (FIG. 3c). Histone locus bodies have high concentrations of histone pre-mRNAs and are often closely associated with Cajal bodies, which are thought to be involved in the production and recycling of certain snRNPs. Experiments of tethering histone pre-mRNA to chromatin revealed the RNA to be sufficient for the formation of these bodies[65]. Similarly, tethering of β-globin mRNA to chromatin results in the formation of speckles[65], which are nuclear structures that contain a high concentration of splicing factors. Together, these data indicate that nascent RNA has a key role in organizing nuclear architecture.

Conclusions and future directions

Evidence is accumulating that nascent RNAs, both coding and non-coding, regulate transcription and chromatin modification. All pre-mRNAs may thus be considered as 'bifunctional RNAs' (REF. 66), possessing both coding and regulatory functions.

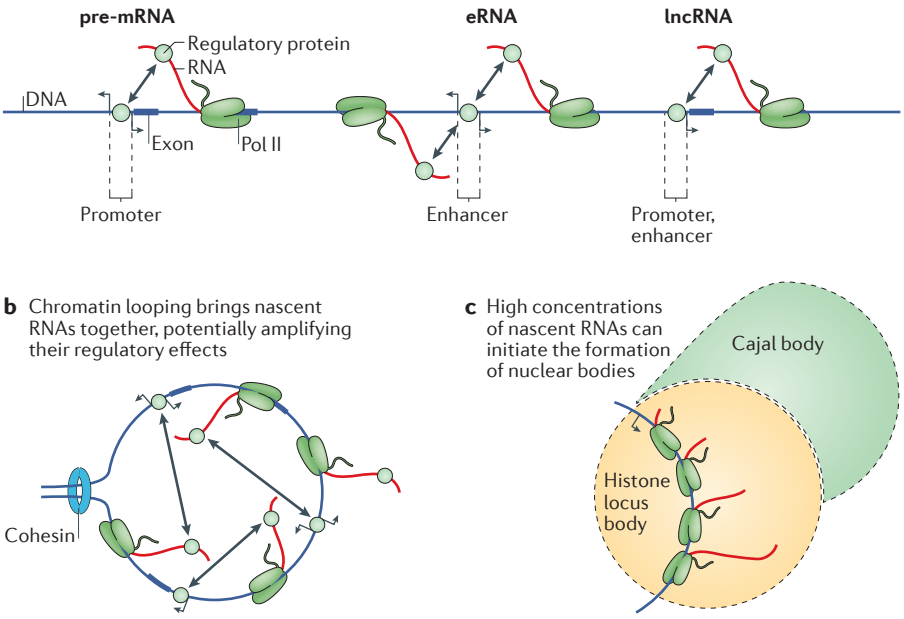

Figure 3 | **Functional similarities between pre-mRNA and nascent ncRNAs and the formation of higher-order chromatin structures. a** | The interaction of transcription and chromatin regulatory proteins (for example, yin and yang 1 (YY1), WD repeat-containing 5 (WDR5) and Polycomb repressive complex 2 (PRC2); green circles) with both pre-mRNAs and nascent non-coding RNAs (ncRNAs), such as enhancer RNAs (eRNAs) and activatory long ncRNAs (lncRNAs), suggests a degree of functional equivalence between these three types of nascent transcripts, which may all regulate the binding of proteins to adjacent regulatory regions of chromatin in *cis*, be these promoters or enhancers. **b** | Regulatory interactions between ncRNAs and protein-coding genes may function in both directions, with the pre-mRNA also having a regulatory impact on enhancer function and on other genes. Furthermore, chromatin looping may amplify the regulatory effects of individual RNAs to wider chromatin loci by bringing together neighbouring genes and regulatory elements. **c** | Pre-mRNA is sufficient to initiate the formation of nuclear bodies, such as histone locus bodies, Cajal bodies and speckles (not shown), suggesting a role in nuclear organization. Pol II, RNA polymerase II.

What is the advantage of nascent RNA having a direct role in transcription and chromatin regulation? Binding to nascent RNA, which is the product of transcription, provides the most direct (and therefore precise) means of feedback to regulate transcription and co-transcriptional processes. It also supports a highly dynamic system, both temporally and spatially. The ability of specific sequence elements to regulate RNA processing, transcription elongation and chromatin modification ensures that these processes remain coordinated and occur at specific positions within the gene. Promiscuous RNA binding by transcription and chromatin regulatory proteins provides a means to detect transcription per se and promote, or antagonize, the recruitment of these factors. This may allow the formation of positive-feedback loops in which active gene expression states can be maintained. Compared with the relatively spatially restricted chromatin fibre, the length and flexibility of nascent RNA increase the nuclear volume over which a particular gene or regulatory element can make contact with soluble proteins (the 'search space').

A number of experimental approaches will be required to establish how general is the function of nascent RNA as a regulator of its own expression. More CLIP experiments are required to confirm direct interactions of regulatory factors with RNAs in living cells and to identify where on the RNAs these interactions take place. When seeking to demonstrate the function of a particular nascent RNA, knockdown experiments are inherently challenging owing to the transient nature of nascent RNAs and the potential lack of access by the RNAi machinery, and genetic mutation introduces the confounding effect of changing the DNA sequence. Antisense oligonucleotides (reviewed in REF. 67)

have been used with varying success to degrade specific nascent RNAs, and CRISPR-mediated RNA cleavage[68] is expected to provide further advances. To determine whether nascent RNA is sufficient to modulate the recruitment of regulatory factors, the ability to tether RNA to particular sites on chromatin will be important[54,65,69]. In parallel, the identification of the amino acids of transcription and chromatin regulators that function in RNA binding will be necessary to establish the importance of RNA interaction for the activity of the proteins. When considering the role of local RNA concentrations, measurements of how RNA is distributed in 3D space will be essential for understanding the effect on chromatin modification and higher-order chromatin structure. The discovery that, like DNA, histones and the Pol II CTD, RNA is also extensively chemically modified[70], will further extend our understanding of the regulatory role of RNA in its nascent form.

Lenka Skalska, Manuel Beltran-Nebot and Richard G. Jenner are at the UCL Cancer Institute, University College London, 72 Huntley Street, London WC1E 6BT, UK.

Jernej Ule is at the Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK; and The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK.

Correspondence to J.U. and R.G.J.
j.ule@ucl.ac.uk;
r.jenner@ucl.ac.uk

doi:10.1038/nrm.2017.12
Published online 8 Mar 2017;
corrected online 9 Mar 2017

1. Peterlin, B. M. & Price, D. H. Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell* **23**, 297–305 (2006).
2. Tsai, M. C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689–693 (2010).
3. Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).
4. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750–756 (2008).
5. Pandey, R. R. *et al.* Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **32**, 232–246 (2008).
6. Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012).
7. Kim, T.-K. & Shiekhattar, R. Diverse regulatory interactions of long noncoding RNAs. *Curr. Opin. Genet. Dev.* **36**, 73–82 (2016).
8. Li, W., Notani, D. & Rosenfeld, M. G. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* **17**, 207–223 (2016).
9. Lee, J. T. Epigenetic regulation by long noncoding RNAs. *Science* **338**, 1435–1439 (2012).
10. Beltran, M. *et al.* The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Res.* **26**, 896–907 (2016).
11. Hendrickson, D. G., Kelley, D. R., Tenen, D., Bernstein, B. & Rinn, J. L. Widespread RNA binding by chromatin-associated proteins. *Genome Biol.* **17**, 28 (2016).
12. Proudfoot, N. J. Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. *Science* **352**, aad9926 (2016).
13. Santos-Pereira, J. M. & Aguilera, A. R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* **16**, 583–597 (2015).
14. Ji, X. *et al.* SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* **153**, 855–868 (2013).
15. Yamaguchi, Y., Inukai, N., Narita, T., Wada, T. & Handa, H. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Mol. Cell. Biol.* **22**, 2918–2927 (2002).
16. Pagano, J. M. *et al.* Defining NELF-E RNA binding in HIV-1 and promoter-proximal pause regions. *PLoS Genet.* **10**, e1004090 (2014).
17. Missra, A. & Gilmour, D. S. Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the *Drosophila* RNA polymerase II transcription elongation complex. *Proc. Natl Acad. Sci. USA* **107**, 11301–11306 (2010).
18. Saldi, T., Cortazar, M. A., Sheridan, R. M. & Bentley, D. L. Coupling of RNA polymerase II transcription elongation with pre-mRNA splicing. *J. Mol. Biol.* **428**, 2623–2635 (2016).
19. Naftelberg, S., Schor, I. E., Ast, G. & Kornblihtt, A. R. Regulation of alternative splicing through coupling with transcription and chromatin structure. *Annu. Rev. Biochem.* **84**, 165–198 (2015).
20. Custódio, N. & Carmo-Fonseca, M. Co-transcriptional splicing and the CTD code. *Crit. Rev. Biochem. Mol. Biol.* **51**, 395–411 (2016).
21. Alexander, R. D., Innocente, S. A., Barrass, J. D. & Beggs, J. D. Splicing-dependent RNA polymerase pausing in yeast. *Mol. Cell* **40**, 582–593 (2010).
22. Chathoth, K. T., Barrass, J. D., Webb, S. & Beggs, J. D. A splicing-dependent transcriptional checkpoint associated with prespliceosome formation. *Mol. Cell* **53**, 779–790 (2014).
23. Nojima, T. *et al.* Mammalian NET-Seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* **161**, 526–540 (2015).
24. Jonkers, I., Kwak, H. & Lis, J. T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* **3**, e02407 (2014).
25. Veloso, A. *et al.* Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res.* **24**, 896–905 (2014).
26. Mayer, A. *et al.* Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **161**, 541–554 (2015).
27. Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953 (2013).
28. de Almeida, S. F. *et al.* Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat. Struct. Mol. Biol.* **18**, 977–983 (2011).
29. Kim, S., Kim, H., Fong, N., Erickson, B. & Bentley, D. L. Pre-mRNA splicing is a determinant of histone H3K36 methylation. *Proc. Natl Acad. Sci. USA* **108**, 13564–13569 (2011).
30. Convertini, P. *et al.* Sudemycin E influences alternative splicing and changes chromatin modifications. *Nucleic Acids Res.* **42**, 4947–4961 (2014).
31. Yuan, W. *et al.* Heterogeneous nuclear ribonucleoprotein L Is a subunit of human KMT3a/Set2 complex required for H3 Lys-36 trimethylation activity *in vivo*. *J. Biol. Chem.* **284**, 15701–15707 (2009).
32. Zhao, J. *et al.* Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* **40**, 939–953 (2010).
33. Davidovich, C., Zheng, L., Goodrich, K. J. & Cech, T. R. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. Mol. Biol.* **20**, 1250–1257 (2013).
34. Kaneko, S., Son, J., Shen, S. S., Reinberg, D. & Bonasio, R. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1258–1264 (2013).
35. Kaneko, S., Son, J., Bonasio, R., Shen, S. S. & Reinberg, D. Nascent RNA interaction keeps PRC2 activity poised and in check. *Genes Dev.* **28**, 1983–1988 (2014).
36. Davidovich, C. *et al.* Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA. *Mol. Cell* **57**, 552–558 (2015).
37. Riising, E. M. *et al.* Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol. Cell* **55**, 347–360 (2014).
38. Jermann, P., Hoerner, L., Burger, L. & Schubeler, D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc. Natl Acad. Sci. USA* **111**, 3415–3421 (2014).
39. Cifuentes-Rojas, C., Hernandez, A. J., Sarma, K. & Lee, J. T. Regulatory interactions between RNA and polycomb repressive complex 2. *Mol. Cell* **55**, 171–185 (2014).
40. Herzog, V. A. *et al.* A strand-specific switch in noncoding transcription switches the function of a Polycomb/Trithorax response element. *Nat. Genet.* **46**, 973–981 (2014).
41. Chen, P. B., Chen, H. V., Acharya, D., Rando, O. J. & Fazzio, T. G. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat. Struct. Mol. Biol.* **22**, 999–1007 (2015).
42. Wei, C. *et al.* RBFox2 binds nascent RNA to globally regulate Polycomb complex 2 targeting in mammalian genomes. *Mol. Cell* **62**, 875–889 (2016).
43. Di Ruscio, A. *et al.* DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature* **503**, 371–376 (2013).
44. Holz-Schietinger, C. & Reich, N. O. RNA modulation of the human DNA methyltransferase 3A. *Nucleic Acids Res.* **40**, 8550–8557 (2012).
45. Savell, K. E. *et al.* Extra-coding RNAs regulate neuronal DNA methylation dynamics. *Nat. Commun.* **7**, 12091 (2016).
46. Tilgner, H. *et al.* Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res.* **22**, 1616–1625 (2012).
47. Fan, J. *et al.* Long non-coding RNA ROR decoys gene-specific histone methylation to promote tumorigenesis. *Genome Biol.* **16**, 139 (2015).
48. Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124 (2011).
49. Lai, F. *et al.* Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature* **494**, 497–501 (2013).
50. Li, W. *et al.* Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013).
51. Yang, L. *et al.* lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature* **500**, 598–602 (2013).
52. Yang, Y. W. *et al.* Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *eLife* **3**, e02046 (2014).
53. Jeon, Y. & Lee, J. T. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* **146**, 119–133 (2011).
54. Sigova, A. A. *et al.* Transcription factor trapping by RNA in gene regulatory elements. *Science* **350**, 978–981 (2015).
55. Kim, T. K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
56. Engreitz, J. M. *et al.* Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).
57. Vance, K. W. & Ponting, C. P. Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.* **30**, 348–355 (2014).
58. Nickerson, J. A., Krochmalnic, G., Wan, K. M. & Penman, S. Chromatin architecture and nuclear RNA. *Proc. Natl Acad. Sci. USA* **86**, 177–181 (1989).
59. Hall, L. L. *et al.* Stable CoT-1 repeat RNA is abundant and associated with euchromatic interphase chromosomes. *Cell* **156**, 907–919 (2014).
60. Cerase, A., Pintacuda, G., Tattermusch, A. & Avner, P. Xist localization and function: new insights from multiple levels. *Genome Biol.* **16**, 166 (2015).
61. Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
62. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
63. Kung, J. T. *et al.* Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell* **57**, 361–375 (2015).

64. Saldaña-Meyer, R. *et al.* CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* **28**, 723–734 (2014).
65. Shevtsov, S. P. & Dundr, M. Nucleation of nuclear bodies by RNA. *Nat. Cell Biol.* **13**, 167–173 (2011).
66. Ulveling, D., Francastel, C. & Hubé, F. When one is better than two: RNA with dual functions. *Biochimie* **93**, 633–644 (2011).
67. Kole, R., Krainer, A. R. & Altman, S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.* **11**, 125–140 (2012).
68. O'Connell, M. R. *et al.* Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516**, 263–266 (2014).
69. Shechner, D. M., Hacisuleyman, E., Younger, S. T. & Rinn, J. L. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* **12**, 664–670 (2015).
70. Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G. & Suzuki, T. RNA modifications: what have we learned and where are we headed? *Nat. Rev. Genet.* **17**, 365–372 (2016).
71. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).
72. Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I. & Chédin, F. GC skew at the 5′ and 3′ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res.* **23**, 1590–1600 (2013).
73. Skourti-Stathaki, K., Kamieniarz-Gdula, K. & Proudfoot, N. J. R-Loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436–439 (2014).
74. Castellano-Pozo, M. *et al.* R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol. Cell* **52**, 583–590 (2013).
75. Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chédin, F. R-Loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* **45**, 814–825 (2012).
76. Boque-Sastre, R. *et al.* Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc. Natl Acad. Sci. USA* **112**, 5785–5790 (2015).
77. Sanz, L. *et al.* Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals. *Mol. Cell* **63**, 167–178 (2016).
78. Wang, X. *et al.* Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. *Nature* **454**, 126–130 (2008).
79. Bose, D. A. *et al.* RNA binding to CBP stimulates histone acetylation and transcription. *Cell* **168**, 135–149 (2017).
80. Zhao, D. Y. *et al.* SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature* **529**, 48–53 (2016).
81. Hatchi, E. *et al.* BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol. Cell* **57**, 636–647 (2015).

Acknowledgements
The authors are grateful to J. Rinn for his comments on the manuscript.

Competing interests statement
The authors declare no competing interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (box)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF

ERRATUM

Regulatory feedback from nascent RNA to chromatin and transcription

Lenka Skalska, Manuel Beltran-Nebot, Jernej Ule and Richard G. Jenner

Nature Reviews Molecular Cell Biology http://dx.doi.org/10.1038/nrm.2017.12 (2017)

In the version of the article originally published online, an incorrect sentence in the legend of Figure 1 has now been corrected. The statement now reads: "The methyltransferase KMT3A is recruited to specific locations within the gene through mechanisms dependent on splicing and interactions with heterogeneous nuclear RNP L (HNRNPL), contributing to the enrichment of histone H3 Lys36 trimethylation (H3K36me3) at exons." In addition, the configuration of arrows in Figure 3b was incorrect. This has now been corrected.