REVIEW

The diverse requirements of ARS2 in nuclear cap-binding complex-dependent RNA processing

Connor O'Sullivan, Perry L. Howard

Department of Biochemistry and Microbiology, University of Victoria, PO Box 1700 STN CSC, Victoria, BC, V8W 2Y2 Canada

Correspondence: Perry L. Howard
E-mail: phoward@uvic.ca
Received: June 17, 2016
Published: August 02, 2017

ARS2 is a stable component of the nuclear cap-binding complex (CBC) and is critical for RNA Polymerase II transcript processing. Moreover, ARS2, and its orthologue SERRATE in plants, has been implicated in having a role in most established CBC-dependent functions. This review will provide insight into the functions of ARS2/SERRATE in numerous RNA Polymerase II transcript processing events, which happen co-transcriptionally from initiation to termination, and post-transcriptionally during maturation and export into the cytoplasm. Additionally, we will discuss what is known regarding ARS2/SERRATE structure in plants and in mammals.

Keywords: ARS2/SERRATE; Cap-binding complex (CBC); RNA Polymerase II; exosome; microRNA; RNA export; replication-dependent histone RNA; neural stem cells; heterochromatin

To cite this article: Connor O'Sullivan, *et al*. The diverse requirements of ARS2 in nuclear cap-binding complex-dependent RNA processing. RNA Dis 2017; 4: e1376. doi: 10.14800/rd.1376.

Introduction

The early life of all RNA polymerase II (RNAP II) transcripts starts with the addition of a 7-methylguanosine (m7G) cap at the 5'-end [1]. The 5'-m7G cap is co-transcriptionally bound by the nuclear cap-binding complex (CBC), which, in addition to protecting the transcript from degradation, interacts with several multi-protein complexes in a mutually exclusive manner to couple the 5'-cap with transcript processing, turnover, and export. RNAP II generates several classes of RNA, including messenger RNA (mRNA), microRNA (miRNA), replication-dependent histone (RDH) mRNA, small nuclear RNA (snRNA), some small nucleolar RNA (snoRNA), and long noncoding RNA (lncRNA). Each of these RNA classes possess unique processing requirements. The interaction between the CBC and processing machinery is necessary for transcription termination, 3'-end formation, exosomal degradation, intranuclear transport, and export into the cytoplasm [2-11]. Precisely how the cap facilitates these disparate processes and discriminates between transcripts is not known. The cap-binding protein, ARS2, is a scaffold that interacts constitutively with CBP20/80 and the 5'-end cap to form a complex called CBCA [8, 10, 11], which mediates interactions between the cap complex and RNA 3'-end processing, shuttling, exosome, and export machinery. Thus, ARS2 is emerging as a critical factor, physically coupling multiple steps in the life of RNAP II transcripts.

ARS2 was first identified as a gene that conferred arsenic

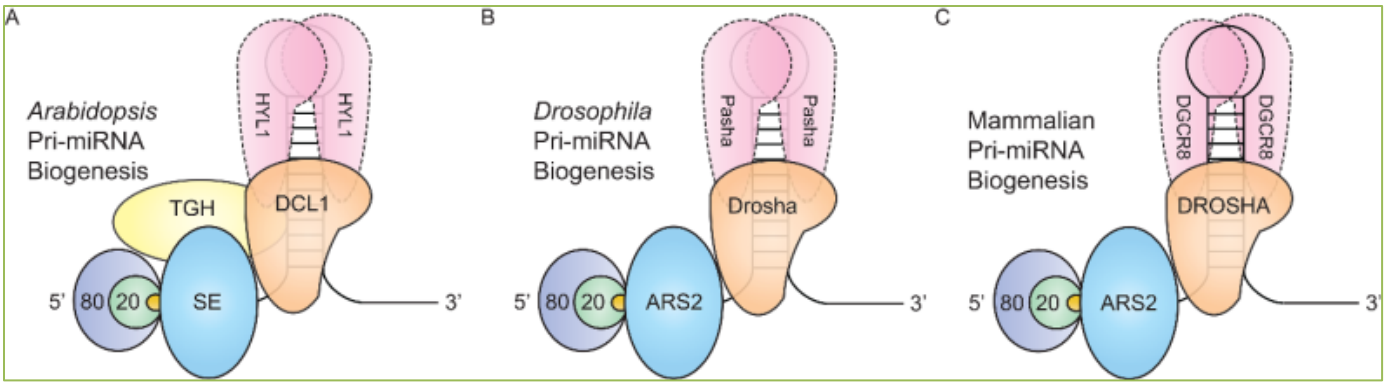

Figure 1. ARS2 and microRNA biogenesis. ARS2/SERRATE has a conserved role in miRNA biogenesis. A) In *Arabidopsis* SERRATE (SE) interacts with CBP80/ABH1 (80) and CBP20 (20) at the 5' m7G cap (small yellow circle) and is required for efficient and accurate pri-miRNA cleavage by DCL1. SE also interacts with HYL1 and TGH. B, C) In *Drosophila* and mammals, ARS2 is required for efficient pri-miRNA cleavage by the Microprocessor, which consists of Drosha and Pasha. C) DGCR8 is the mammalian orthologue of Pasha. In mammals, ARS2 interacts with DROSHA and is required for efficient processing and miRNA stability.

resistance to cells [12]. This was later shown to be due to expression of a truncated cDNA that resulted in a dominant-negative phenotype; instead, ARS2 expression correlates with arsenic sensitivity [8]. Early genetic studies determined the requirements of ARS2, or its plant orthologue SERRATE (SE), during development. ARS2/SERRATE knockout results in embryonic lethality in plants, fission yeast, fruit flies, zebrafish, and mice [13-18]. Conditional *Ars2* knockout in hematopoietic tissues results in decreased cellularity in the bone marrow [8], while conditional knockout in mouse subventricular zone neural stem cells (NSCs) decreased neurogenic and self-renewal capacity [19]. Conversely, high levels of ARS2 in NSCs correlates with increased neurogenesis and an elongated life span in mice [20]. Additionally, there are several recent reports of ARS2 dysregulation in human cancers [21-23]. Together, these findings indicate ARS2 has a vital role in the development and maintenance of diverse tissues and suggest that maintaining appropriate levels of ARS2 may be important for human disease.

It is increasingly evident that ARS2 function is intimately intertwined with cap-binding proteins CBP20/80. Immunoprecipitation of either ARS2 or CBP20/80 pulls down a substantially overlapping set of proteins, implying a strong functional overlap [10, 11]. This is supported by loss-of-function experiments demonstrating ARS2 and CBP20/80 regulate similar transcripts, and phenotypically resemble one another [10, 11]. This review will emphasize the interactions mediated by ARS2, for additional reviews on the CBC please see [24, 25]. We will discuss the early functional studies linking ARS2/SERRATE to the CBC and miRNA and RDH biogenesis, review what is known about ARS2 domain structure and function, as well as the recent "omic" data implicating ARS2 in linking transcription to the surveillance, export, and silencing machineries. The theme emerging is that ARS2 is involved in the regulation of most capped RNAP II transcripts. ARS2 acts as a scaffold protein able to bind both RNA and proteins and provide a physical link between the 5'-CBC and the protein complexes involved in 3'-end processing, maturation, degradation, export, and silencing.

ARS2/SERRATE and microRNA biogenesis

The first studies to characterize ARS2 function were in *Arabidopsis* with its orthologue SERRATE (SE). Hypomorphic mutations of SE in *Arabidopsis* result in pleiotropic developmental defects [26-28], which had overlapping phenotypes with mutants of genes involved in miRNA biogenesis, as well as with CBP80/ABH1 and CBP20 mutants [13, 27-31]. For reviews of miRNA biogenesis in *Arabidopsis* please refer to [32, 33]. Briefly, in *Arabidopsis*, pri-miRNAs are cleaved to miRNA in the nucleus by a complex composed of the RNase III enzyme DICER-LIKE 1 (DCL1), the double-stranded (ds)RNA binding protein HYPONASTIC LEAVES 1 (HYL1), and the scaffold TOUGH (TGH) (Figure 1) [32]. Following DCL1 cleavage, the miRNA/miRNA* duplex is methylated by HEN1, which protects the miRNA from degradation, and is exported into the cytoplasm by HASTY [34-36]. Once exported, miRNA are bound by ARGONAUTE 1 (AGO1) and induce degradation or inhibit translation of its target transcripts [37, 38].

Early studies showed that SE, within the context of the CBC, is required for proper miRNA biogenesis [13, 30]. SE physically interacts with pri-miRNA, as well as with the CBC [39-41]. Both *cbp* and *se* mutants accumulate levels of pri-miRNA [7, 42], indicating their involvement at the early stages of processing. Indeed, SE and HYL1 directly interact with DCL1 and are required for the efficient and accurate cleavage of pri-miRNA and pre-miRNA [40, 43-46]. SE also

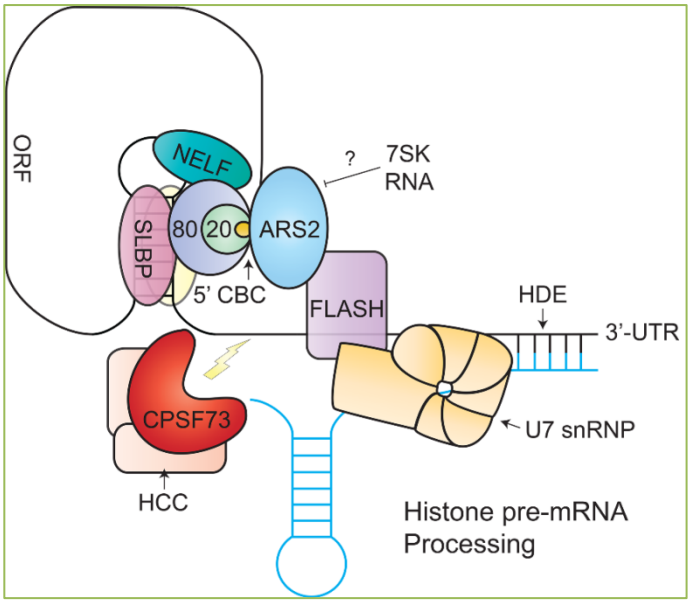

Figure 2. ARS2 and replication-dependent histone processing. Binding of the U7 snRNP to the HDE and FLASH are required for endonucleolytic cleavage by CPSF73, which is part of a histone pre-mRNA cleavage complex (HCC). CBCA makes multiple contacts with the 3' processing machinery, which are required for accurate cleavage. CBP80 interacts with SLBP and NELF, and ARS2 interacts with FLASH. Additionally, histone 3'-end formation is negatively regulated by the noncoding 7SK RNA through an interaction with ARS2.

interacts with the scaffold TGH [47]. HEN1 interacts with the same regions of DCL1 and HYL1 as SE, and does not interact with SE itself, supporting the notion that SE acts upstream of miRNA methylation and is possibly released in order for HEN1 to function [48].

The role of ARS2/SERRATE and the CBC in miRNA biogenesis is conserved in metazoans [8, 49]. In mammals, RNAP II generates capped primary miRNA (pri-miRNA) transcripts [50, 51], which are trimmed in the nucleus by the Microprocessor complex consisting of the RNase III DROSHA and the RNA binding DiGeorge syndrome chromosomal region 8 (DGCR8) protein [52, 53]. ARS2 interacts with DROSHA and is required for the stability and efficient processing of pri-miRNA to precursor miRNA (pre-miRNA) by the Microprocessor [8]. Interestingly, ARS2 is thought to act as a cofactor for DROSHA, as it influences the efficiency and specificity of DROSHA activity [8, 9, 54].

In *Drosophila*, the miRNA biogenesis process is very similar; pri-miRNAs are cleaved to pre-miRNA in the nucleus by the Microprocessor complex [55]. Following export into the cytoplasm, pre-miRNA are cleaved by Dicer-1 to generate mature miRNA, which are then loaded onto an Ago1-dependent RNA-induced silencing complex (RISC) [56-58]. As in mammals, *Drosophila* Ars2 (dArs2) interacts with the CBC and Microprocessor, and is required for pri-miRNA processing and stability [49]. *Drosophila* also use RNAi for innate immunity to protect against viral infection [49]. Viral dsRNA is processed into small interfering RNA (siRNA) by Dicer-2 in the cytoplasm [56, 59]. Dicer-2, along with the dsRNA-binding protein R2D2, are required for loading Ago2-RISC, which mediates siRNA silencing [56, 59, 60]. Interestingly, dARS2 also interacts with cytoplasmic Dicer-2 and is required for the processing of presumably uncapped long dsRNA into siRNA [49]. Depletion of dArs2, CBP20, or CBP80 in flies results in an increased susceptibility to infections by RNA viruses, due to defective siRNA biogenesis [49]. Thus, ARS2, CBP20, and CBP80, may have cofactor roles in RNA processing that extend beyond their role in the nuclear cap complex. Collectively, this work established that ARS2/SE has a conserved role in miRNA biogenesis, and physically couples the CBC to the Microprocessor in insects and animals, or to DCL1 in plants (Figure 1).

ARS2 and replication-dependent histone RNA

The demonstration that ARS2/SERRATE is part of the CBC that physically bridges the cap to the Microprocessor, raised the possibility that ARS2 would be required for other CBC-dependent processes. The first indication of this came from work by Kiriyama *et al.* and Gruber *et al.* who showed that ARS2 is required for processing of replication-dependent histone (RDH) transcripts [9, 61]. As cells enter S phase, histone proteins are rapidly synthesized to package newly replicated DNA [62]. RDH transcripts are transcribed by RNAP II, m7G capped, and intronless [63]. At their 3'-ends, RDH transcripts contain a conserved stem-loop, and are the only known metazoan mRNA not polyadenylated [63]. Instead, RDH transcripts undergo endonucleolytic cleavage, via cleavage and polyadenylation specificity factor (CPSF) subunit CPSF73, between the conserved 3'-end stem-loop and a histone downstream element (HDE) [64-66]. RDH processing is cap-dependent and requires the coordination of several multimeric complexes. For example, the stem-loop is bound by stem-loop binding protein (SLBP), while the HDE base pairs with U7 snRNA, which is part of a multi-subunit U7 small nuclear ribonucleoprotein (snRNP) complex that acts as a molecular ruler to guide endonucleolytic cleavage by the histone pre-mRNA cleavage complex (HCC), which consists of CPSF73, CPSF100, and symplekin [67-72]. In addition, negative elongation factor (NELF) not only associates with histone loci at the promoter, but is also required for processing RDH pre-mRNA at the 3'-end (Figure 2) [6]. Misprocessing of RDH transcripts results in aberrant read-through and polyadenylation [9, 73].

ARS2 mediates interactions with RDH pre-mRNA and components of the RDH 3'-processing machinery. CBP20/80

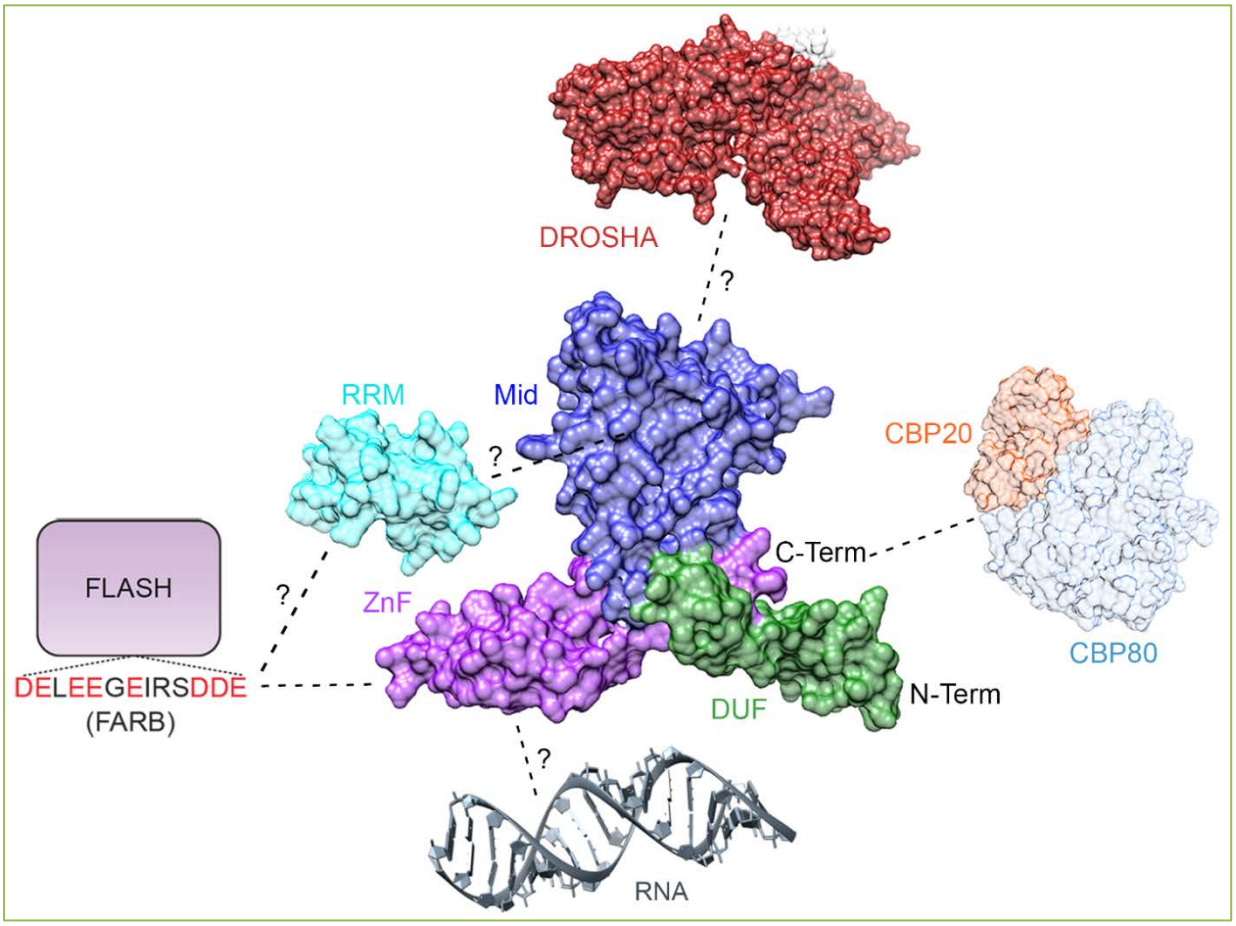

Figure 3. Mammalian ARS2 structure/function. A representation of mouse ARS2 domains (center) along with their reported interacting partners, based on [54]. The structured part of the protein consists of the domain of unknown function (DUF3546) shown in green, Mid domain in dark blue, and ZnF domain in purple. The orientation of the RRM (shown in cyan) relative to the core is unknown. The unstructured C-terminus (C-term) is required for interaction with CBP80/20, the Mid domain has been implicated as being important for interaction with DROSHA, the ZnF is required for RNA interaction, and FLASH through the FLASH ARS2 binding (FARB) peptide interacts with the ZnF and RRM. How ARS2 precisely contacts RNA/protein is not understood. CBP80/20 structure corresponds to PDB 1N52 [155], DROSHA to PDB 5B16 [53], and the RNA structure is from PDB 5DV7 [156].

interacts with SLBP and negative elongation factor (NELF) [6, 11]. Meanwhile, mammalian ARS2 interacts with RDH RNA and FLICE-associated huge protein (FLASH) [9, 54, 61]. In turn, FLASH directly interacts with the U7 snRNP component LSM11, and is required for RDH pre-mRNA endonucleolytic cleavage [74-77]. Knockdown (KD) of either ARS2 or FLASH disrupts the formation of histone locus bodies (HLBs), increases aberrant polyadenylated histone transcripts, and decreases histone protein levels [9, 54, 61, 78]. Most likely as a result, cells deficient in either ARS2 or FLASH are delayed in S-phase progression [54, 61, 78]. As an added layer of complexity, the noncoding 7SK RNA, an important factor in regulating transcriptional elongation, interacts with ARS2 and negatively regulates RDH 3'-end processing, potentially by sequestering ARS2 [9]. Further work is needed to understand how 7SK RNA regulates ARS2, as well as the precise role of ARS2 in RDH 3'-end formation (Figure 2). Nevertheless, the finding that ARS2, as part of the CBC, interacts with both miRNA and RDH processing machinery suggested ARS2 has a broader role in physically coupling RNAP II transcript processing.

ARS2/SERRATE structure/function

The next important milestone in understanding ARS2 function was obtaining the SE crystal structure by the Yuan Lab [46]. Their work showed that SE consists of 3 domains in a walking man-like conformation, with the domain of unknown function (DUF3546), Mid, and zinc finger (ZnF) domains forming the leading leg, body, and lagging leg, respectively; the N and C-termini are unstructured [46]. Using homology modeling, bioinformatics analysis and this critical structural information, our lab determined that the DUF3546, Mid and ZnF domains were conserved between plants and

metazoans [54]. Additionally, we identified a previously uncharacterized RNA recognition motif (RRM) that is highly conserved within metazoans but absent in plants [54]. This information provided the basis for deletion and point mutant analysis to assess the functions of these domains in *Arabidopsis* and mammalian cells, respectively.

The interactions between SE, DCL1, and RNA are primarily mediated through the unstructured arginine-rich N-terminus, ZnF domain and unstructured C-terminus [40, 46]. *In vitro* binding assays have mapped the RNA binding sites to a strong affinity site located within the unstructured N-terminus and a weaker site located within the ZnF domain and unstructured C-terminus [40, 46]. Similarly, DCL1 interacts both with the unstructured N-terminus and the ZnF domain [40]. Interestingly, the unstructured N-terminus is dispensable for pri-miRNA processing *in vitro* [40], suggesting the interactions through the ZnF and unstructured C-terminus are sufficient. In support of this hypothesis, the SE ZnF is essential for stimulating pri-miRNA cleavage *in vitro* [40]. In fact, the ZnF along with the C-terminal tail are sufficient to rescue the *se-1* mutant morphology and miRNA levels *in vivo* [46]. Thus, the ZnF and C-terminus form a critical core for miRNA processing, with additional non-essential interactions formed through the N-terminus.

A somewhat different picture is emerging for mammalian ARS2 interactions. Point mutant analysis of mouse ARS2 (mARS2) has revealed the ZnF domain is important for both protein and RNA interactions [54]. Mutations in the ZnF domain abolished interactions with both pri-miRNA and RDH pre-mRNA transcripts in a pull-down assay, and disrupted miRNA and RDH processing based on reporter assays [54]. The interaction between the ZnF and these RNA appears to be direct, as wildtype mARS2, but not the ZnF mutants, was able to associate with RNA under high salt wash conditions, which removed mARS2 associated proteins [54]. The ZnF domain is also required for interacting with FLASH through the FLASH ARS2 binding (FARB) peptide [54], indicating this domain can mediate both protein-protein and protein-RNA interactions (Figure 3).

An important distinguishing feature of metazoan ARS2 is the presence of an RRM domain, which is inserted into an unstructured loop region between helix 4 and 5 of the Mid domain, and is predicted to fold similar to splicing factor 3b subunit 4 (SF3B4) [54]. Mutation analysis showed the ARS2 RRM is important for histone expression and interactions with FLASH [54]. However, mutations of the RNP1 and RNP2 motifs within the domain did not affect binding to miRNA or RDH mRNA, suggesting this domain is not involved in RNA interactions [54]. It remains possible that the RRM may mediate interaction with other types of RNA, bind RNA through other regions of the domain not mutated, or exclusively interact with proteins, similar to Y14 and UPF3 RRMs (Figure 3) [79, 80]. Further investigation is warranted to distinguish between these possibilities.

The ARS2 DUF3546 domain is defined by three tyrosine residues that are conserved in yeast, plants, and metazoans, suggesting they are functionally important [54]. Mutating the conserved tyrosines to phenylalanine in mARS2 abolished interaction with pri-miRNA and RDH pre-mRNA in low salt conditions; however, the DUF mutant could bind RNA as efficiently as wildtype in high salt conditions [54]. These results point to an allosteric mechanism or an indirect effect through a RNA-binding cofactor at this site. Interestingly, published mass spectrometry data show ARS2 can be phosphorylated at Y172 [81] and Y175 [82] within the DUF3546 domain, suggesting the function of this domain may be regulated post-translationally by phosphorylation. In fission yeast, mutation of F165 in the DUF3546 domain contributes to a temperature sensitive allele [83]. How the DUF3546 domain regulates interactions with RNA requires further investigation.

Mapping the interaction between mARS2 and DROSHA has proven to be difficult. In contrast to SE, mutations to the ZnF or deletion of the C-terminus did not affect binding to DROSHA [54]. Machida *et al*. originally proposed the idea that the Mid domain of SE/ARS2 may function as a platform for protein-protein interactions [46]. Consistent with this hypothesis, mutations in two separate regions of the mARS2 Mid domain partially disrupted DROSHA pulldown and affected miRNA biogenesis as evidenced by reporter assays [54], thus implicating the Mid domain in DROSHA binding (Figure 3). However, the distance between the Mid domain mutation sites suggest these mutations likely locally disrupt the folding of this domain, and thus the exact binding site has not been elucidated.

ARS2 interacts with the 5'-CBC through the unstructured proline-rich C-terminus [10, 54]. Using a yeast two-hybrid assay, Hallais *et al*. mapped the CBC interaction site to amino acids 502-871 of the C-terminus [10]. Using deletion analysis, our lab found that amino acids 763-875 were necessary to immunoprecipitate the CBC with ARS2 [54]. Interestingly, this region was also necessary for interactions with histone mRNA, as deleting this region abolished the ability of ARS2 to co-precipitate histone mRNA, but not miRNA [54]. This suggests that the ZnF domain alone is not sufficient for ARS2 to interact with RDH mRNA, and additional interactions are required, likely mediated through the C-terminus of ARS2 and the CBC. Collectively, this work established that ARS2, through its domains and unstructured termini, is able to interact with the CBC,

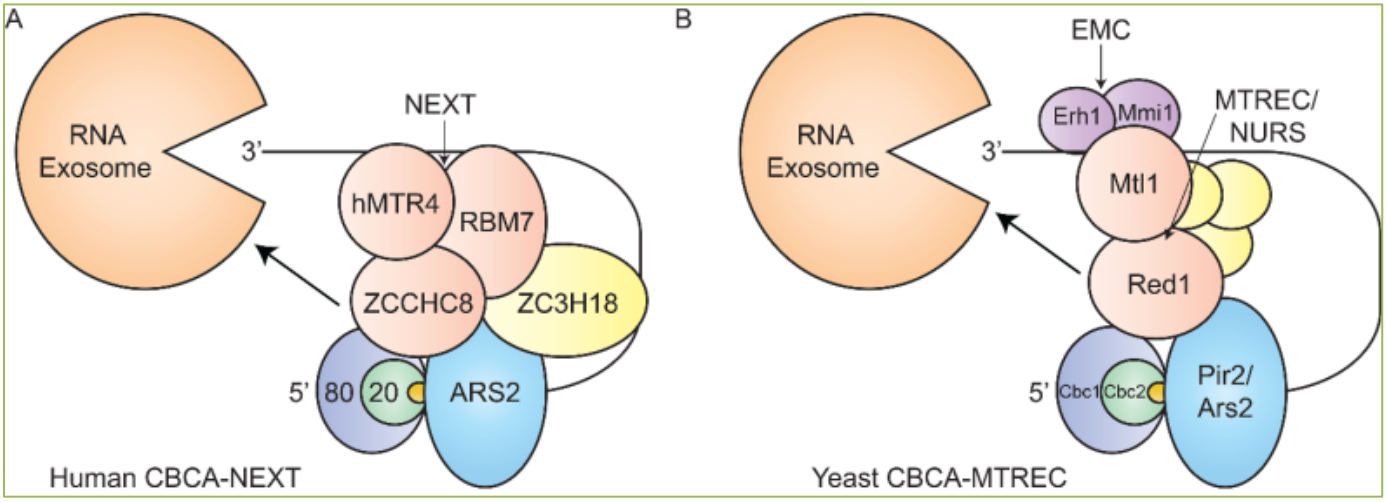

Figure 4. ARS2 and the exosome. A) CBC-ARS2 targets aberrant transcripts to the RNA exosome through the NEXT complex, composed of hMTR4, ZCCHC8 and RBM7, and the NEXT-associated protein ZC3H18. B) *S. pombe* Pir2/Ars2 and CBC proteins Cbc1/Cbc2 target aberrant transcripts to the exosome through the MTREC/NURS complex, minimally composed of Mtl1 and Red1, and associated with several subcomplexes, including the Erh1-Mmi1 complex (EMC) (see ARS2 and heterochromatin).

different types of RNA, and different unique processing machineries. Thus, ARS2 has the properties of a scaffold able to mediate multiple protein and RNA interactions, which positions it as an ideal candidate linking the cap complex to a broader role in RNAP II transcript processing.

ARS2 and transcription termination

Confirmation of a broader role in RNAP II transcription came with the demonstration that ARS2 and the CBC are required for inducing cap-proximal transcription termination for snRNA, RDH RNA, promoter upstream transcripts (PROMPTs) and mRNA. This was demonstrated through proteomic analysis using immunoprecipitations of the machinery involved in these processes, and knockdown experiments of ARS2 or CBP20/80, which increases 3'-read-through transcripts for each of these RNA classes [10, 11]. Interestingly, transcripts longer than ~1 kb were largely unaffected by ARS2 depletion with regards to 3'-read-through [10], suggesting CBCA or a CBCA-interacting factor have a mechanism to limit their activity to promoter proximal areas. However, very little is currently known mechanistically about how ARS2 may interact with these processes.

ARS2 and the exosome

A major role of the CBC is in RNA quality control, and in limiting the effects of promiscuous RNAP II transcription. Recently, ARS2 was shown to participate in targeting transcripts to the nuclear RNA exosome through interaction with the human nuclear exosome targeting (NEXT) complex [11, 84]. This complex recruits the exosome to degrade PROMPTs and 3'-extended snRNA [11, 84-87]. NEXT is composed of the RNA helicase hMTR4, the zinc-knuckle ZCCHC8, and the RRM-containing RBM7 [11, 84]. Affinity capture mass spectrometry revealed a stoichiometric interaction between CBCA, NEXT, and another ZnF protein ZC3H18 [11]. Depletion of CBCA, NEXT components, or ZC3H18, results in accumulation of PROMPTs [10, 11]. RBM7 associates with newly synthesized RNA, is enriched at regions close to the 5'-cap, and RBM7-RNA interaction is disrupted following CBC KD [85]. These data indicate that ARS2 and the CBC recruit the NEXT complex to newly synthesized RNAP II transcripts that are destined for exosomal destruction (Figure 4).

The role of ARS2 and the cap complex in RNA quality control is conserved in fission yeast. In *Schizosaccharomyces pombe* (*S. pombe*), Pir2/Ars2 (yeast orthologue) and Cbc1-Cbc2 (orthologues of CBP80 and CBP20, respectively) interact with the Mtl1-Red1 core (MTREC) complex [88,89] (alternatively named nuclear RNA silencing (NURS) complex) [90]. Mtl1 is a Mtr4-like helicase and Red1 is a ZnF protein [88, 91]. Despite not having sequence similarities, CBCA-MTREC is likely the fission yeast functional equivalent of the human CBCA-NEXT complex. MTREC, along with Pir2/Ars2 and the CBC, are essential for targeting RNA to the exosome in *S. pombe* (Figure 4) [88-91]. Similar to the fate of PROMPTs in mammals, MTREC in *S. pombe* delivers polyadenylated cryptic unstable transcripts (CUTs) to the exosome for destruction [89].

ARS2 and mRNA 3'-end processing

3'-end processing of mRNA consists of endonucleolytic

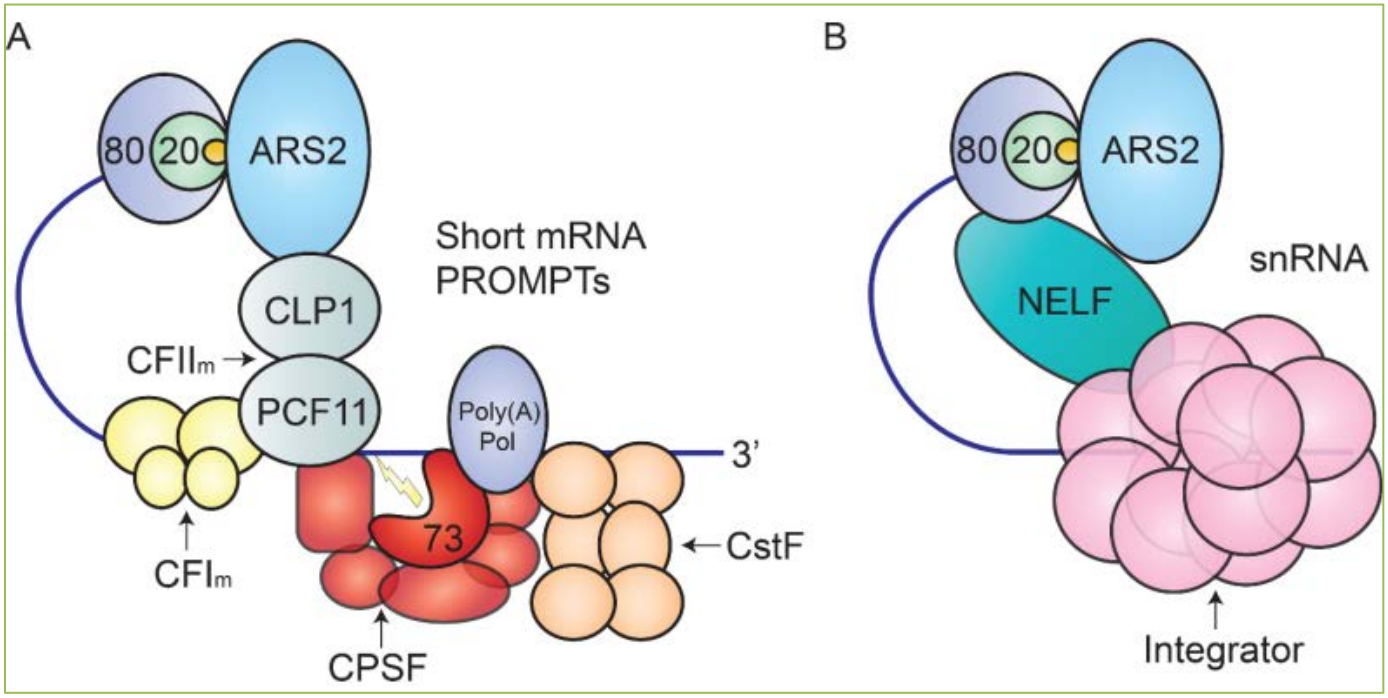

Figure 5. ARS2 and 3'-end formation. A) ARS2 is required for 3'-end cleavage by CPSF73 at short mRNA transcripts and PROMPTs, and interacts with CLP1/PCF11 of the $CFII_m$ complex, which bridges the CPSF and CFI_m complexes. The CstF complex and Poly(A) polymerase are also shown. B) ARS2 is also required for 3'-end formation by the Integrator complex at snRNA transcripts. CBCA may mediate an interaction with Integrator through NELF.

cleavage followed by polyadenylation. The cleavage and polyadenylation machinery are recruited co-transcriptionally through the CTD of RNAP II [92]. A large number of factors are required for proper mRNA processing. These include the multi-subunit CPSF, cleavage stimulation factor (CstF), cleavage factor I_m (CFI_m) and $CFII_m$ complexes as well as poly(A) polymerase [93, 94]. $CFII_m$ bridges RNAP II and the nascent transcript to CPSF and CFI_m [95], and is required for pre-mRNA cleavage by CPSF73 prior to polyadenylation [93]. CBCA interacts with and stabilizes the pre-mRNA 3'-processing machinery, and is required for efficient cleavage, but not polyadenylation of short pre-mRNA [5,10]. Interestingly, ARS2 was shown to interact with $CFII_m$ component CLP1, and depletion of either ARS2, CLP1, or another $CFII_m$ component PCF11, induces 3'-read-through reminiscent of CBCA KD, suggesting the interaction is functional (Figure 5A) [10, 96].

ARS2 and snRNA 3'-end processing

ARS2 is also required for snRNA 3'-end formation [10], a process which shares NELF with RDH 3'-end processing machinery [6, 97]. snRNA 3'-end formation is mediated by the twelve subunit Integrator complex that contains homologues of CPSF73 and CPSF100 (Int11 and Int9, respectively) [98]. NELF interacts with Integrator and is required for accurate snRNA 3'-processing [97]. Thus, similar to its role in RDH 3'-end formation, CBCA may be affecting snRNA processing through an interaction with NELF (Figure 5B).

ARS2 and export

In metazoans, export of snRNA requires CBCA, phosphorylated adapter for RNA export (PHAX), the export receptor CRM1/XPO1, and RanGTP [99, 100]. PHAX directly interacts with the CBC and snRNA [10, 11, 100, 101]. Notably, ARS2 stimulates PHAX binding to the CBC, promoting the formation of a stable complex called CBCAP, composed of CBC, ARS2, and PHAX [10]. The mechanism underlying the allosteric regulation of PHAX binding by ARS2 is currently unclear and requires further investigation. However, PHAX is phosphorylated by CK2 kinase, and PHAX must be in its phosphorylated state in order for CRM1 to be recruited along with RanGTP, and for snRNA export to occur (Figure 6A) [100, 102]. In addition to snRNAs, the CBCAP complex can bind m7G capped snoRNAs, and PHAX binding is required for their intranuclear transport to Cajal bodies, where snoRNAs are further processed [10,103]. This suggests ARS2 may also be involved in snoRNA transport.

The mechanism of the specific interaction between PHAX and snRNA provides insight into understanding how different types of transcripts may be distinguished [10]. Differential processing of snRNAs is achieved through

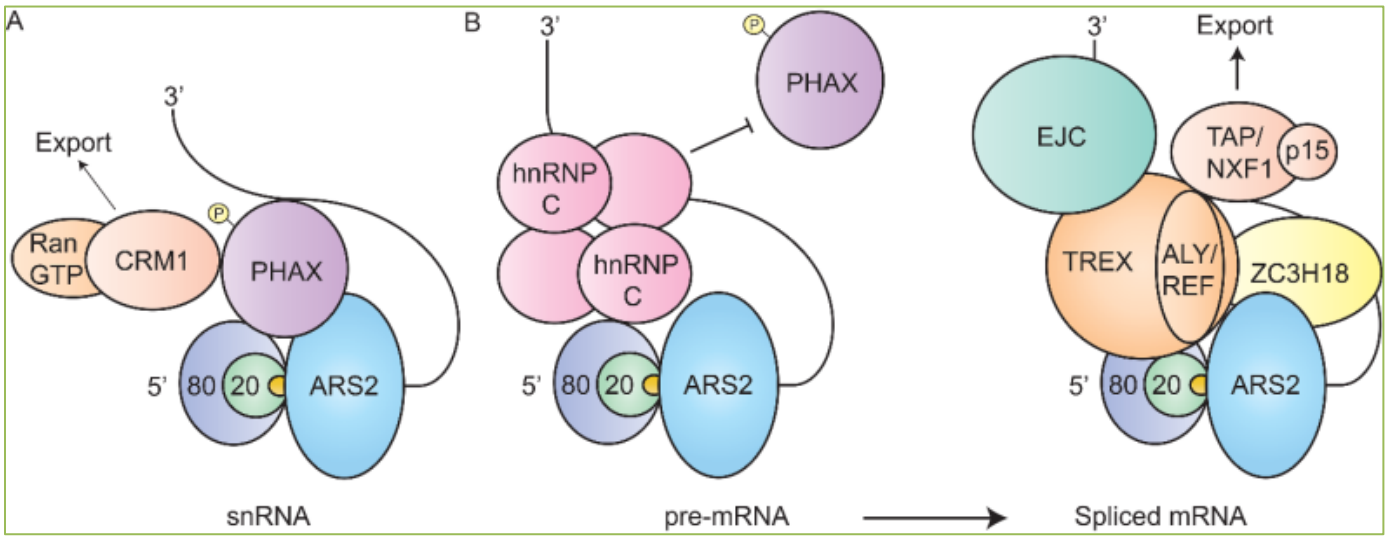

Figure 6. ARS2 and export. A) ARS2 mediates snRNA export by stimulating phosphorylated adaptor for RNA export (PHAX) binding to CBC, forming the CBCAP complex. snRNA bound to CBCAP is then exported by CRM1/RanGTP. B) Transcripts longer than ~200-300 nt are bound by the hnRNP C tetramer, which interacts with the CBC and inhibits PHAX binding, thereby committing them to the mRNA export pathway (left). Splicing and the presence of CBCA, ZC3H18, and the exon junction complex (EJC) stimulates TREX recruitment to the 5'-end of mRNA. The export adaptor ALY/REF, a component of TREX, mediates handover to TAP/NXF1 and p15 for export into the cytoplasm (right).

preferential binding of the heterogeneous nuclear ribonucleoprotein (hnRNP) C tetramer to transcripts longer than 200-300nt [104]. hnRNP C directly interacts with CBP80 and RNA, and competitively inhibits PHAX binding [104]. Therefore, U snRNAs, which are typically <200 nt, bind PHAX by default and are exported via CRM1-RanGTP, while longer mRNA transcripts are inhibited from using this pathway through the action of hnRNP C (Figure 6B) [104, 105].

CBCA also plays an important role in mRNA export, largely mediated by the multisubunit transcription export (TREX) complex, which acts as an adaptor for the export receptor TAP/NXF1 [106-111]. Human TREX is composed of the THO complex (THOC1, THOC2, THOC5, THOC6, THOC7, Tex1), CIP29, UAP56 and ALY/REF [112-114]. TREX is involved in release of mRNA from nuclear speckle domains and coupling to the export receptor [115]. The presence of the exon junction complex (EJC), which is deposited following splicing, stimulates TREX recruitment [106, 107, 110, 112, 116], ensuring correctly processed transcripts are exported. Following binding of TREX subunits to mRNA, ALY/REF and THOC5 handover the mRNA to the TAP/NXF1-p15 heterodimer [117], which then transports the mRNA through the nuclear pore [111, 118].

ALY/REF directly interacts with CBP20/80 [104, 119, 120], while THOC2, UAP56, and CIP29 independently interact with ARS2 [113], and these interactions are required for efficient mRNA export [120, 121]. ARS2 KD results in accumulation of mRNA in nuclear speckle domains [121]. Interestingly, Zinc-knuckle protein ZC3H18, which interacts with ARS2 as part of the CBC-NEXT complex [11, 84, 85], also interacts with the TREX complex [121]. ZC3H18 KD prevents efficient TREX recruitment to RNA and also results in an accumulation of mRNA in nuclear speckle domains [121]. Although the details are less understood, intronless mRNA also relies on components of this pathway, including CBCA, TREX, ZC3H18, and TAP/NXF1 [108, 109, 120, 121]. Thus, CBCA, through ZC3H18, may control recruitment of TREX and export of transcripts that are capped and correctly processed (Figure 6C).

Recently, an alternative mammalian CBC was discovered where NCBP3 could bind CBP80 and the m7G cap in place of CBP20 [122]. Interestingly, both NCBP3 and CBP20 bound many common factors, including ARS2, and only double KD of NCBP3 and CBP20 significantly disrupted mRNA export, suggesting some redundancy between the two complexes; however, NCBP3 preferentially interacted with TREX, while CBP20 exclusively bound snRNA and PHAX, indicating these two complexes may have developed specialized functions in mRNA and snRNA export [122]. Deciphering the roles of this alternative CBC as well as how it interacts with ARS2 will be an exciting area of research.

ARS2/SERRATE and splicing

Pre-mRNA splicing, whereby introns are removed and exons are ligated together in a two-step transesterification reaction, is carried out by the spliceosome complex. For excellent, detailed reviews on splicing, see [123-125]. In *Arabidopsis*, CBP80/ABH1, CBP20 and SE are required for

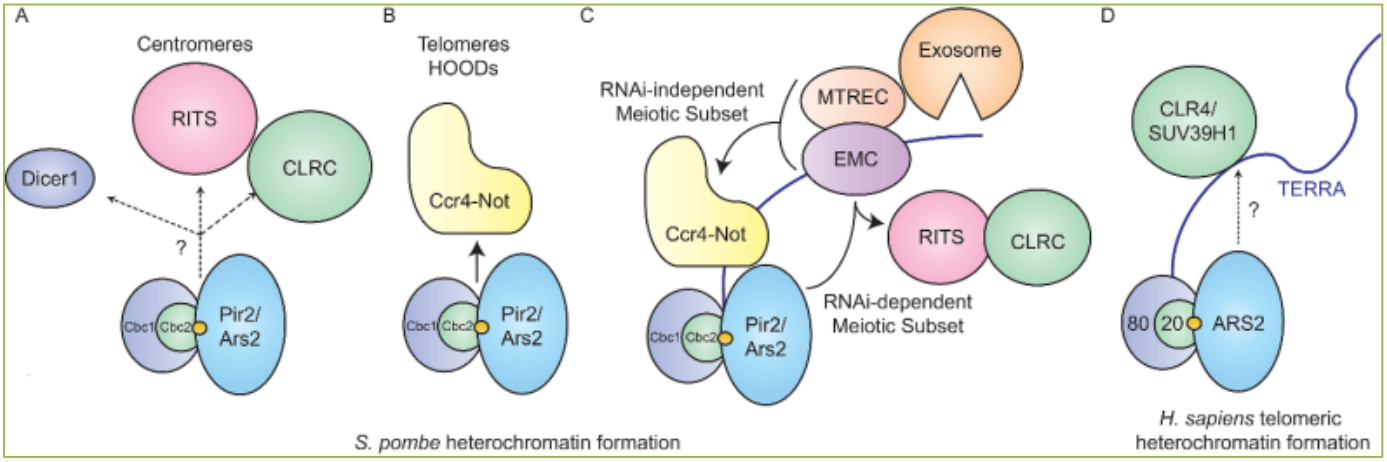

Figure 7. ARS2 and heterochromatin. A) *S. pombe* Pir2/Ars2 is required for heterochromatin formation at centromeres. Centromeric heterochromatin formation requires RNAi machinery (Dicer1 and RITS) as well as CLRC. Whether Pir2/Ars2 mediates centromeric heterochromatin through Dicer1, RITS and/or CLRC requires further study. B) Pir2/Ars2 also regulates heterochromatin formation at telomeres and heterochromatin domains (HOODs) in a Ccr4-Not dependent manner. C) Pir2/Ars2 regulates heterochromatin at a subset of meiotic loci through RNAi-independent and RNAi-dependent mechanisms. The RNAi-independent subset is thought to be mediated by the EMC through its ability to recruit MTREC and Ccr4-Not. The RNAi-dependent subset requires RITS and CRLC, which are thought to be recruited through the EMC. D) Human ARS2 regulates telomeric repeat-containing RNA (TERRA), which interact with the CLR4/SUV39H1 methyltransferase and are correlated with H3K9me levels. The mechanism of TERRA regulation by ARS2 and its involvement in telomeric heterochromatin formation is not known (indicated by question mark).

cap-proximal splicing and alternative splicing, as mutants primarily affect retention of the first intron and alternative 5'-splice site selection [7, 41, 126]. The mechanistic details of how SE mediates splicing have not been resolved. In mammals, the CBC facilitates cap-proximal splicing, interacts with the U4/U6·U5 tri-snRNP in a RNA-independent manner, is required for co-transcriptional spliceosome assembly, and is involved in alternative splicing [127-130]. Although mammalian ARS2 co-purifies with multiple splicing factors [11, 131, 132], whether ARS2 directly participates in splicing in metazoans has not been established.

ARS2 and heterochromatin formation

Recently, a requirement for ARS2 in heterochromatin formation has been shown in fission yeast. In *S. pombe*, Pir2/Ars2 is required for heterochromatin formation at centromeres, telomeres, heterochromatin domains (HOODs), and a subset of meiotic loci [83, 90]. Pir2/Ars2 KD results in decreased histone H3 lysine 9 methylation (H3K9me) levels at these sites, a hallmark of heterochromatin formation [83,90]. The requirement of Pir2/Ars2 for heterochromatin formation at these diverse regions involves several partially overlapping pathways. The best characterized pathway is centromeric silencing, which is mediated by RNA interference (RNAi) machinery and heterochromatin formation [133], and requires Pir2/Ars2 [83, 90]. Transcription from repeat elements in these regions forms dsRNAs, which are cleaved by Dicer to generate siRNAs that are loaded onto an Ago1-containing RNA-induced transcriptional silencing (RITS) complex [134]. RITS then targets the complex to peri-centromeric regions for silencing and heterochromatin formation [134]. RITS is composed of Ago1, the chromodomain protein Chp1, and the GW protein Tas3 [134,135]. Heterochromatin formation is mediated through an association between the RITS complex and the Clr4-Rik1-Cul4 (CLRC) complex, comprised of Clr4, a histone methyltransferase, Rik1, a heterochromatin targeting protein, and Cul4, an E3 ubiquitin ligase [136]. Together, RITS and the CLRC complex are responsible for the initiation of H3K9me [137]. Clr4 then binds to H3K9me and promotes the spread of heterochromatin [137]. Given the ability of dArs2 to associate with Dicer2 in *Drosophila*, and dArs2's role in siRNA biogenesis in this organism [49], it is possible that Pir2/Ars2 is regulating heterochromatin at the centromeres through interactions with Dicer. However, Pir2/Ars2 also associates with RITS-associated CLRC complex in pull-down assays, and therefore likely has additional roles in silencing beyond enhancing Dicer activity (Figure 7A) [90].

The second Pir2/Ars2-dependent pathway for heterochromatin formation is through the Ccr4-Not complex. This complex, along with Pir2/Ars2, mediates heterochromatin formation at telomeres, HOODs, and a subset of meiotic loci [83]. Ccr4-Not in *S. pombe* is comprised of Ccr4, Caf1, Caf40, and five Not proteins (Not1-5) [138]. Ccr4-Not has deadenylation (Ccr4 and Caf1) and ubiquitination (Not4) activity and is responsible for mRNA

turnover in the cytoplasm [138]. Interestingly, both enzymatic activities of the complex are required for heterochromatin formation, although the mechanisms behind these requirements is currently unclear [139]. Pir2/Ars2 co-purifies with all seven subunits of the Ccr4-Not complex, and heterochromatin formation at telomeres, HOODS, and some meiotic islands is Pir2/Ars2/Ccr4-Not-dependent (Figure 7B) [83]. A similar role for dArs2 and CCR4-NOT was found in retrotransposon silencing at telomeres in *Drosophila* [140, 141].

Meiotic loci in fission yeast represent a third type of heterochromatin silencing. Meiotic transcripts contain a determinant of selective removal (DSR) sequence in their 3'-UTR, recognized during the vegetative state by Mmi1, part of the Erh1-Mmi1 complex (EMC), which targets them for degradation via the exosome (Figure 7) [83, 88, 90, 91, 142]. The EMC is also required for heterochromatin formation at meiotic loci during the vegetative state through both RNAi-independent and RNAi-dependent mechanisms [83]. The RNAi-independent process is facilitated through the ability of the EMC to recruit MTREC and the exosome [83, 88-91, 143], and is thought to rely on the ability of the MTREC complex to independently interact with the Ccr4-Not complex to mediate H3K9me [83, 138, 139]. An alternative mechanism of meiotic loci silencing is RNAi-dependent and is mediated through the EMC's ability to recruit the RITS complex and associated CLRC [83, 144]. Both RNAi-dependent and -independent processes rely on Pir2/Ars2 (Figure 7C) [83,90], but it is currently unclear how Pir2/Ars2 is restricted to a subset of meiotic loci.

In humans, ARS2 has been implicated in telomeric heterochromatin formation through its regulation of long noncoding telomeric repeat-containing RNA (TERRA) [145]. TERRA are transcribed by RNAP II, associate with telomeric DNA, and have been implicated in telomere maintenance [146]. ARS2 co-purifies with TERRA, and ARS2 KD in HeLa cells increases TERRA levels and the abundance of TERRA associated with telomeres [145]. Furthermore, TERRA itself directly interacts with the human homolog of the yeast Clr4 histone methyltransferase (SUV39H1) [147], and H3K9me levels are correlated with TERRA levels [147-149]. In *Arabidopsis*, SE and CBP80/CBP20 regulate the levels of hundreds of long noncoding RNAs (lncRNAs) [150]. Thus, the role of ARS2 in lncRNA biogenesis is conserved, and ARS2 is likely important for the function of other lncRNAs in mammals. Taken together, ARS2 is tied to heterochromatin formation at diverse regions using multiple pathways in yeast and metazoans (Figure 7D).

An anomalous role for ARS2 as a transcription factor

The vast majority of ARS2 functions involve its interactions with CBP20/80. However, there is one report of ARS2 as a transcription factor that does not fit this model. As mentioned, ARS2 has a key role in maintaining NSCs in the mouse brain [19]. Conditional *Ars2* knockout in NSCs decreased their self-renewal capacity and multipotency, with their fate skewed towards astroglial production [19]. This phenotype was rescued by SOX2, a transcription factor essential for NSC maintenance [19]. Curiously, ARS2 bound to a small region within the SOX2 enhancer in a RNA-independent but cell type-dependent manner [19]. The presence of this enhancer region was necessary for SOX2 luciferase reporter expression following ARS2 overexpression [19]. Furthermore, expression of SOX2 was sufficient to rescue the defects in NSC self-renewal and multipotency of ARS2-deficient animals [19]. This work implicated ARS2 as a transcription factor. However, how ARS2 regulates SOX2 transcription is currently unclear. It is complicated by the fact that the *Sox2* gene is located within an intron of the lncRNA SOX2OT (overlapping transcript) [151], and SOX2OT positively regulates SOX2 [152-154]. As mentioned, ARS2 is implicated in lncRNA function. Thus, it is plausible that some of the effects of ARS2 on SOX2 expression may be related to ARS2 regulation of SOX2OT biogenesis. Further work is needed to discern the precise role of ARS2 in SOX2 expression.

Concluding Remarks

RNAP II transcription, from initiation to termination and export, is a highly coordinated multistep process that requires the CBC, yet there is little information on how this coordination is mediated. ARS2/SE is emerging as a critical component of the nuclear CBC, able to form a physical link between the CBC and numerous protein complexes that regulate different steps in RNAP II transcript biogenesis, turnover, and export. In yeast, these processes are also coupled to epigenetic chromatin modifications. Whether this coupling of RNA turnover to heterochromatin formation is conserved in metazoans remains an open question. However, the recent discovery of the role of ARS2 in *Drosophila* germline telomeric repeat silencing and TERRA regulation in humans suggests similar mechanisms, potentially lncRNA-dependent, will occur in metazoans. Understanding how ARS2 functions in these diverse processes will provide insight into the mechanisms of how these diverse processes are functionally coupled.

Conflicting interests

The authors have declared that no conflict of interests exist.

Acknowledgements

PLH is supported through a Discovery Grant from the National Science and Engineering Research Council (NSERC) of Canada.

Author Contributions

C.O. and P.L.H. wrote the manuscript.

Abbreviations

m7G: 7-methylguanosine; AGO: Argonaute; CBC: cap-binding complex; $CF1_m$, cleavage factor I_m; $CFII_m$: cleavage factor II_m; CPSF: cleavage and polyadenylation specificity factor; CstF: cleavage stimulation factor; CLRC: Clr4-Rik1-Cul4; CUTs: cryptic unstable transcripts; DSR: determinant of selective removal; DCL1: DICER-LIKE1; DGCR8: DiGeorge syndrome chromosomal region 8; DUF: domain of unknown function; dsRNA: double-stranded RNA; Erh1: Enhancer of rudimentary; EMC: Erh1-Mmi1 complex; FARB: FLASH ARS2 binding peptide; FLASH: FLICE-associated huge protein; HOODs, heterochromatin domains; hnRNP: heterogeneous ribonucleoprotein; HDE: histone downstream element; H3K9me: histone H3 lysine 9 methylation; HLBs: histone locus bodies; HCC: histone pre-mRNA cleavage complex; HYL1: HYPONASTIC LEAVES 1; KD: knockdown; lncRNA, long noncoding RNA; mRNA: messenger RNA; miRNA: microRNA; MTREC: Mtl1-Red1 core; NELF: negative elongation factor; NSCs: neural stem cells; NEXT: nuclear RNA exosome targeting; NURS: nuclear RNA silencing; PHAX: phosphorylated adaptor for RNA export; pA: polyadenylation; pre-miRNA: precursor miRNA; pri-miRNA: primary miRNA; PROMPTs: promoter upstream transcripts; RDH: replication-dependent histone; RISC: RNA-induced silencing complex; RITS: RNA-induced transcriptional silencing; RNAi: RNA interference; RNAP II: RNA polymerase II; CTD: RNAP II C-terminal domain; RRM: RNA recognition motif; *S. pombe*: *Schizosaccharomyces pombe*; SE: SERRATE; siRNA: small interfering RNA; snRNA: small nuclear RNA; snRNP: small nuclear ribonucleoprotein; snoRNA: small nucleolar RNA; SF3B4: splicing factor 3b subunit 4; SOX2OT: SOX2 overlapping transcript; SLBP: stem loop binding protein; TERRA: telomeric repeat-containing RNA; TGH: TOUGH; TREX: transcription export; ZnF: zinc finger.

References

1. Lewis JD, Izaurralde E. The role of the cap structure in RNA processing and nuclear export. Eur J Biochem FEBS 1997; 247:461-469.
2. Furuichi Y, LaFiandra A, Shatkin AJ. 5'-Terminal structure and mRNA stability. Nature 1977; 266:235-239.
3. Izaurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW. A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell 1994; 78:657-668.
4. Visa N, Izaurralde E, Ferreira J, Daneholt B, Mattaj IW. A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. J Cell Biol 1996; 133:5-14.
5. Flaherty SM, Fortes P, Izaurralde E, Mattaj IW, Gilmartin GM. Participation of the nuclear cap binding complex in pre-mRNA 3' processing. Proc Natl Acad Sci U S A 1997; 94:11893-11898.
6. Narita T, Yung TMC, Yamamoto J, Tsuboi Y, Tanabe H, Tanaka K, *et al*. NELF interacts with CBC and participates in 3' end processing of replication-dependent histone mRNAs. Mol Cell 2007; 26:349-365.
7. Laubinger S, Sachsenberg T, Zeller G, Busch W, Lohmann JU, Rätsch G, *et al*. Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. Proc Natl Acad Sci U S A 2008; 105:8795-8800.
8. Gruber JJ, Zatechka DS, Sabin LR, Yong J, Lum JJ, Kong M, *et al*. Ars2 links the nuclear cap-binding complex to RNA interference and cell proliferation. Cell 2009; 138:328-339.
9. Gruber JJ, Olejniczak SH, Yong J, La Rocca G, Dreyfuss G, Thompson CB. Ars2 promotes proper replication-dependent histone mRNA 3' end formation. Mol Cell 2012; 45:87-98.
10. Hallais M, Pontvianne F, Andersen PR, Clerici M, Lener D, Benbahouche NEH, *et al*. CBC-ARS2 stimulates 3'-end maturation of multiple RNA families and favors cap-proximal processing. Nat Struct Mol Biol 2013; 20:1358-1366.
11. Andersen PR, Domanski M, Kristiansen MS, Storvall H, Ntini E, Verheggen C, *et al*. The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat Struct Mol Biol 2013; 20:1367-1376.
12. Rossman TG, Wang Z. Expression cloning for arsenite-resistance resulted in isolation of tumor-suppressor fau cDNA: possible involvement of the ubiquitin system in arsenic carcinogenesis. Carcinogenesis 1999; 20:311-316.
13. Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J. SERRATE: a new player on the plant microRNA scene. EMBO Rep 2006; 7:1052-1058.
14. Oh SW, Kingsley T, Shin HH, Zheng Z, Chen HW, Chen X, *et al*. A P-element insertion screen identified mutations in 455 novel essential genes in Drosophila. Genetics 2003; 163:195-201.
15. Golling G, Amsterdam A, Sun Z, Antonelli M, Maldonado E, Chen W, *et al*. Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. Nat Genet 2002; 31:135-140.
16. Amsterdam A, Nissen RM, Sun Z, Swindell EC, Farrington S, Hopkins N. Identification of 315 genes essential for early zebrafish development. Proc Natl Acad Sci U S A 2004; 101:12792-12797.
17. Wilson MD, Wang D, Wagner R, Breyssens H, Gertsenstein M, Lobe C, *et al*. ARS2 is a conserved eukaryotic gene essential for early mammalian development. Mol Cell Biol 2008; 28:1503-1514.

18. Kim DU, Hayles J, Kim D, Wood V, Park HO, Won M, *et al*. Analysis of a genome-wide set of gene deletions in the fission yeast Schizosaccharomyces pombe. Nat Biotechnol 2010; 28:617-623.

19. Andreu-Agullo C, Maurin T, Thompson CB, Lai EC. Ars2 maintains neural stem-cell identity through direct transcriptional activation of Sox2. Nature 2012; 481:195-198.

20. Carrasco-Garcia E, Arrizabalaga O, Serrano M, Lovell-Badge R, Matheu A. Increased gene dosage of Ink4/Arf and p53 delays age-associated central nervous system functional decline. Aging Cell 2015; 14:710-714.

21. He Q, Cai L, Shuai L, Li D, Wang C, Liu Y, *et al*. Ars2 is overexpressed in human cholangiocarcinomas and its depletion increases PTEN and PDCD4 by decreasing microRNA-21. Mol Carcinog 2013; 52:286-296.

22. He Q, Huang Y, Cai L, Zhang S, Zhang C. Expression and prognostic value of Ars2 in hepatocellular carcinoma. Int J Clin Oncol 2014; 19:880-888.

23. Cui L, Gao C, Zhang RD, Jiao Y, Li WJ, Zhao XX, *et al*. Low expressions of ARS2 and CASP8AP2 predict relapse and poor prognosis in pediatric acute lymphoblastic leukemia patients treated on China CCLG-ALL 2008 protocol. Leuk Res 2015; 39:115-123.

24. Topisirovic I, Svitkin YV, Sonenberg N, Shatkin AJ. Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip Rev RNA 2011; 2:277-298.

25. Gonatopoulos-Pournatzis T, Cowling VH. Cap-binding complex (CBC). Biochem J 2014; 457:231-242.

26. Clarke JH, Tack D, Findlay K, Van Montagu M, Van Lijsebettens M. The SERRATE locus controls the formation of the early juvenile leaves and phase length in Arabidopsis. Plant J 1999; 20:493-501.

27. Prigge MJ, Wagner DR. The arabidopsis serrate gene encodes a zinc-finger protein required for normal shoot development. Plant Cell 2001; 13:1263-1279.

28. Grigg SP, Canales C, Hay A, Tsiantis M. SERRATE coordinates shoot meristem function and leaf axial patterning in Arabidopsis. Nature 2005; 437:1022-1026.

29. Bezerra IC, Michaels SD, Schomburg FM, Amasino RM. Lesions in the mRNA cap-binding gene ABA HYPERSENSITIVE 1 suppress FRIGIDA-mediated delayed flowering in Arabidopsis. Plant J 2004; 40:112-119.

30. Yang L, Liu Z, Lu F, Dong A, Huang H. SERRATE is a novel nuclear regulator in primary microRNA processing in Arabidopsis. Plant J 2006; 47:841-850.

31. Papp I, Mur LA, Dalmadi A, Dulai S, Koncz C. A mutation in the Cap Binding Protein 20 gene confers drought tolerance to Arabidopsis. Plant Mol Biol 2004; 55:679-686.

32. Ren G, Yu B. Post-transcriptional control of miRNA abundance in Arabidopsis. Plant Signal Behav 2012; 7:1443-1446.

33. Borges F, Martienssen RA. The expanding world of small RNAs in plants. Nat Rev Mol Cell Biol 2015; 16:727-741.

34. Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, *et al*. Methylation as a Crucial Step in Plant microRNA Biogenesis. Science 2005; 307:932-935.

35. Li J, Yang Z, Yu B, Liu J, Chen X. Methylation Protects miRNAs and siRNAs from a 3′-End Uridylation Activity in Arabidopsis. Curr Biol 2005; 15:1501-1507.

36. Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. Nuclear processing and export of microRNAs in Arabidopsis. Proc Natl Acad Sci U S A 2005; 102:3691-3696.

37. Baumberger N, Baulcombe DC. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci U S A 2005; 102:11928-11933.

38. Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, *et al*. MicroRNAs Inhibit the Translation of Target mRNAs on the Endoplasmic Reticulum in Arabidopsis. Cell 2013; 153:562-574.

39. Wang L, Song X, Gu L, Li X, Cao S, Chu C, *et al*. NOT2 Proteins Promote Polymerase II-Dependent Transcription and Interact with Multiple MicroRNA Biogenesis Factors in Arabidopsis. Plant Cell 2013; 25:715-727.

40. Iwata Y, Takahashi M, Fedoroff NV, Hamdan SM. Dissecting the interactions of SERRATE with RNA and DICER-LIKE 1 in Arabidopsis microRNA precursor processing. Nucleic Acids Res 2013; 41:9129-9140.

41. Raczynska KD, Stepien A, Kierzkowski D, Kalak M, Bajczyk M, McNicol J, *et al*. The SERRATE protein is involved in alternative splicing in Arabidopsis thaliana. Nucleic Acids Res 2014; 42:1224-1244.

42. Kim S, Yang JY, Xu J, Jang IC, Prigge MJ, Chua NH. Two Cap-Binding Proteins CBP20 and CBP80 are Involved in Processing Primary MicroRNAs. Plant Cell Physiol 2008; 49:1634-1644.

43. Fang Y, Spector DL. Identification of Nuclear Dicing Bodies Containing Proteins for MicroRNA Biogenesis in Living Arabidopsis Plants. Curr Biol 2007; 17:818-823.

44. Fujioka Y, Utsumi M, Ohba Y, Watanabe Y. Location of a Possible miRNA Processing Site in SmD3/SmB Nuclear Bodies in Arabidopsis. Plant Cell Physiol 2007; 48:1243-1253.

45. Dong Z, Han MH, Fedoroff N. The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. Proc Natl Acad Sci U S A 2008; 105:9970-9975.

46. Machida S, Chen HY, Yuan YA. Molecular insights into miRNA processing by Arabidopsis thaliana SERRATE. Nucleic Acids Res 2011; 39:7828-7836.

47. Ren G, Xie M, Dou Y, Zhang S, Zhang C, Yu B. Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. Proc Natl Acad Sci U S A 2012; 109:12817-12821.

48. Baranauskė S, Mickutė M, Plotnikova A, Finke A, Venclovas Č, Klimašauskas S, *et al*. Functional mapping of the plant small RNA methyltransferase: HEN1 physically interacts with HYL1 and DICER-LIKE 1 proteins. Nucleic Acids Res 2015; 43:2802-2812.

49. Sabin LR, Zhou R, Gruber JJ, Lukinova N, Bambina S, Berman A, *et al*. Ars2 regulates both miRNA- and siRNA- dependent silencing and suppresses RNA virus infection in Drosophila. Cell 2009; 138:340-351.

50. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, *et al*. MicroRNA genes are transcribed by RNA polymerase II. EMBO J 2004; 23:4051-4060.

51. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are

processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA 2004; 10:1957-1966.

52. Gurung B, Muhammad AB, Hua X. Menin Is Required for Optimal Processing of the MicroRNA let-7a. J Biol Chem 2014; 289:9902-9908.

53. Kwon SC, Nguyen TA, Choi YG, Jo MH, Hohng S, Kim VN, *et al*. Structure of Human DROSHA. Cell 2016; 164:81-90.

54. O'Sullivan C, Christie J, Pienaar M, Gambling J, Nickerson PEB, Alford SC, *et al*. Mutagenesis of ARS2 Domains To Assess Possible Roles in Cell Cycle Progression and MicroRNA and Replication-Dependent Histone mRNA Biogenesis. Mol Cell Biol 2015; 35:3753-3767.

55. Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. Nature 2004; 432:231-235.

56. Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, *et al*. Distinct Roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways. Cell 2004; 117:69-81.

57. Tsutsumi A, Kawamata T, Izumi N, Seitz H, Tomari Y. Recognition of the pre-miRNA structure by Drosophila Dicer-1. Nat Struct Mol Biol 2011; 18:1153-1158.

58. Okamura K, Ishizuka A, Siomi H, Siomi MC. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev 2004; 18:1655-1666.

59. Liu Q, Rand TA, Kalidas S, Du F, Kim HE, Smith DP, *et al*. R2D2, a Bridge Between the Initiation and Effector Steps of the Drosophila RNAi Pathway. Science 2003; 301:1921-1925.

60. Iwasaki S, Sasaki HM, Sakaguchi Y, Suzuki T, Tadakuma H, Tomari Y. Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. Nature 2015; 521:533-536.

61. Kiriyama M, Kobayashi Y, Saito M, Ishikawa F, Yonehara S. Interaction of FLASH with arsenite resistance protein 2 is involved in cell cycle progression at S phase. Mol Cell Biol 2009; 29:4729-4741.

62. DeLisle AJ, Graves RA, Marzluff WF, Johnson LF. Regulation of histone mRNA production and stability in serum-stimulated mouse 3T6 fibroblasts. Mol Cell Biol 1983; 3:1920-1929.

63. Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat Rev Genet 2008; 9:843-854.

64. Ryan K, Calvo O, Manley JL. Evidence that polyadenylation factor CPSF-73 is the mRNA 3′ processing endonuclease. RNA 2004; 10:565-573.

65. Dominski Z, Yang XC, Marzluff WF. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. Cell 2005; 123:37-48.

66. Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, *et al*. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. Nature 2006; 444:953-956.

67. Wang ZF, Whitfield ML, Ingledue TC, Dominski Z, Marzluff WF. The protein that binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. Genes Dev 1996; 10:3028-3040.

68. Dominski Z, Yang XC, Kaygun H, Dadlez M, Marzluff WF. A 3' exonuclease that specifically interacts with the 3' end of histone mRNA. Mol Cell 2003; 12:295-305.

69. Tan D, Marzluff WF, Dominski Z, Tong L. Structure of Histone mRNA Stem-Loop, Human Stem-Loop Binding Protein, and 3′hExo Ternary Complex. Science 2013; 339:318-321.

70. Scharl EC, Steitz JA. The site of 3' end formation of histone messenger RNA is a fixed distance from the downstream element recognized by the U7 snRNP. EMBO J 1994; 13:2432-2440.

71. Yang XC, Torres MP, Marzluff WF, Dominski Z. Three proteins of the U7-specific Sm ring function as the molecular ruler to determine the site of 3'-end processing in mammalian histone pre-mRNA. Mol Cell Biol 2009; 29:4045-4056.

72. Sabath I, Skrajna A, Yang XC, Dadlez M, Marzluff WF, Dominski Z. 3′-End processing of histone pre-mRNAs in Drosophila: U7 snRNP is associated with FLASH and polyadenylation factors. RNA 2013; 19:1726-1744.

73. Sullivan E, Santiago C, Parker ED, Dominski Z, Yang X, Lanzotti DJ, *et al*. Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. Genes Dev 2001; 15:173-187.

74. Yang XC, Burch BD, Yan Y, Marzluff WF, Dominski Z. FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3' end processing of histone pre-mRNAs. Mol Cell 2009; 36:267-278.

75. Burch BD, Godfrey AC, Gasdaska PY, Salzler HR, Duronio RJ, Marzluff WF, *et al*. Interaction between FLASH and Lsm11 is essential for histone pre-mRNA processing in vivo in Drosophila. RNA 2011; 17:1132-1147.

76. Yang XC, Xu B, Sabath I, Kunduru L, Burch BD, Marzluff WF, *et al*. FLASH is required for the endonucleolytic cleavage of histone pre-mRNAs but is dispensable for the 5' exonucleolytic degradation of the downstream cleavage product. Mol Cell Biol 2011; 31:1492-1502.

77. Yang XC, Sabath I, Dębski J, Kaus-Drobek M, Dadlez M, Marzluff WF, *et al*. A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3'-end processing. Mol Cell Biol 2013; 33:28-37.

78. Barcaroli D, Bongiorno-Borbone L, Terrinoni A, Hofmann TG, Rossi M, Knight RA, *et al*. FLASH is required for histone transcription and S-phase progression. Proc Natl Acad Sci U S A 2006; 103:14808-14812.

79. Fribourg S, Gatfield D, Izaurralde E, Conti E. A novel mode of RBD-protein recognition in the Y14-Mago complex. Nat Struct Biol 2003; 10:433-439.

80. Kadlec J, Izaurralde E, Cusack S. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. Nat Struct Mol Biol 2004; 11:330-337.

81. Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, *et al*. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res 2012; 40:D261-270.

82. Amanchy R, Zhong J, Molina H, Chaerkady R, Iwahori A, Kalume DE, *et al*. Identification of c-Src tyrosine kinase substrates using mass spectrometry and peptide microarrays. J Proteome Res 2008; 7:3900-3910.

83. Sugiyama T, Thillainadesan G, Chalamcharla VR, Meng Z, Balachandran V, Dhakshnamoorthy J, *et al*. Enhancer of Rudimentary Cooperates with Conserved RNA-Processing Factors to Promote Meiotic mRNA Decay and Facultative Heterochromatin Assembly. Mol Cell 2016; 61:747-759.

84. Lubas M, Christensen MS, Kristiansen MS, Domanski M, Falkenby LG, Lykke-Andersen S, *et al*. Interaction Profiling Identifies the Human Nuclear Exosome Targeting Complex. Mol Cell 2011; 43:624-637.

85. Lubas M, Andersen PR, Schein A, Dziembowski A, Kudla G, Jensen TH. The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. Cell Rep 2015; 10:178-192.

86. Chlebowski A, Lubas M, Jensen TH, Dziembowski A. RNA decay machines: The exosome. Biochim Biophys Acta BBA - Gene Regul Mech 2013; 1829:552-560.

87. Schneider C, Tollervey D. Threading the barrel of the RNA exosome. Trends Biochem Sci 2013; 38:485-493.

88. Lee NN, Chalamcharla VR, Reyes-Turcu F, Mehta S, Zofall M, Balachandran V, *et al*. Mtr4-like protein coordinates nuclear RNA processing for heterochromatin assembly and for telomere maintenance. Cell 2013; 155:1061-1074.

89. Zhou Y, Zhu J, Schermann G, Ohle C, Bendrin K, Sugioka-Sugiyama R, *et al*. The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome. Nat Commun 2015; 6:7050.

90. Egan ED, Braun CR, Gygi SP, Moazed D. Post-transcriptional regulation of meiotic genes by a nuclear RNA silencing complex. RNA 2014; 20:867-881.

91. Sugiyama T, Sugioka-Sugiyama R. Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. EMBO J 2011; 30:1027-1039.

92. McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, *et al*. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature 1997; 385:357-361.

93. Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates III JR, *et al*. Molecular Architecture of the Human Pre-mRNA 3′ Processing Complex. Mol Cell 2009; 33:365-376.

94. Xiang K, Tong L, Manley JL. Delineating the Structural Blueprint of the Pre-mRNA 3′-End Processing Machinery. Mol Cell Biol 2014; 34:1894-1910.

95. de Vries H, Rüegsegger U, Hübner W, Friedlein A, Langen H, Keller W. Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. EMBO J 2000; 19:5895-5904.

96. Ntini E, Järvelin AI, Bornholdt J, Chen Y, Boyd M, Jørgensen M, *et al*. Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. Nat Struct Mol Biol 2013; 20:923-928.

97. Yamamoto J, Hagiwara Y, Chiba K, Isobe T, Narita T, Handa H, *et al*. DSIF and NELF interact with Integrator to specify the correct post-transcriptional fate of snRNA genes. Nat Commun 2014; 5:4263.

98. Baillat D, Hakimi MA, Näär AM, Shilatifard A, Cooch N, Shiekhattar R. Integrator, a Multiprotein Mediator of Small Nuclear RNA Processing, Associates with the C-Terminal Repeat of RNA Polymerase II. Cell 2005; 123:265-276.

99. Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj IW. A cap-binding protein complex mediating U snRNA export. Nature 1995; 376:709-712.

100. Ohno M, Segref A, Bachi A, Wilm M, Mattaj IW. PHAX, a Mediator of U snRNA Nuclear Export Whose Activity Is Regulated by Phosphorylation. Cell 2000; 101:187-198.

101. Mourão A, Varrot A, Mackereth CD, Cusack S, Sattler M. Structure and RNA recognition by the snRNA and snoRNA transport factor PHAX. RNA 2010; 16:1205-1216.

102. Kitao S, Segref A, Kast J, Wilm M, Mattaj IW, Ohno M. A Compartmentalized Phosphorylation/Dephosphorylation System That Regulates U snRNA Export from the Nucleus. Mol Cell Biol 2008; 28:487-497.

103. Boulon S, Verheggen C, Jady BE, Girard C, Pescia C, Paul C, *et al*. PHAX and CRM1 Are Required Sequentially to Transport U3 snoRNA to Nucleoli. Mol Cell 2004; 16:777-787.

104. McCloskey A, Taniguchi I, Shinmyozu K, Ohno M. hnRNP C Tetramer Measures RNA Length to Classify RNA Polymerase II Transcripts for Export. Science 2012; 335:1643-1646.

105. Masuyama K, Taniguchi I, Kataoka N, Ohno M. RNA length defines RNA export pathway. Genes Dev 2004; 18:2074-2085.

106. Zhou Z, Luo M, Straesser K, Katahira J, Hurt E, Reed R. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature 2000; 407:401-405.

107. Luo MJ, Zhou Z, Magni K, Christoforides C, Rappsilber J, Mann M, *et al*. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. Nature 2001; 413:644-647.

108. Rodrigues JP, Rode M, Gatfield D, Blencowe BJ, Carmo-Fonseca M, Izaurralde E. REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. Proc Natl Acad Sci U S A 2001; 98:1030-1035.

109. Lei H, Dias AP, Reed R. Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. Proc Natl Acad Sci U S A 2011; 108:17985-17990.

110. Gromadzka AM, Steckelberg AL, Singh KK, Hofmann K, Gehring NH. A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs. Nucleic Acids Res 2016; 44:2348-2361.

111. Braun IC, Herold A, Rode M, Conti E, Izaurralde E. Overexpression of TAP/p15 Heterodimers Bypasses Nuclear Retention and Stimulates Nuclear mRNA Export. J Biol Chem 2001; 276:20536-20543.

112. Masuda S, Das R, Cheng H, Hurt E, Dorman N, Reed R. Recruitment of the human TREX complex to mRNA during splicing. Genes Dev 2005; 19:1512-1517.

113. Dufu K, Livingstone MJ, Seebacher J, Gygi SP, Wilson SA, Reed R. ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. Genes Dev 2010; 24:2043-2053.

114. Chi B, Wang Q, Wu G, Tan M, Wang L, Shi M, *et al*. Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. Nucleic Acids Res 2013; 41:1294-1306.

115. Dias AP, Dufu K, Lei H, Reed R. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. Nat Commun 2010; 1:97.

116. Le Hir H, Gatfield D, Izaurralde E, Moore MJ. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J 2001; 20:4987-4997.

117. Hautbergue GM, Hung ML, Golovanov AP, Lian LY, Wilson SA. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. Proc Natl Acad Sci U S A 2008; 105:5154-5159.

118. Fribourg S, Braun IC, Izaurralde E, Conti E. Structural Basis for the Recognition of a Nucleoporin FG Repeat by the NTF2-like Domain of the TAP/p15 mRNA Nuclear Export Factor. Mol Cell 2001; 8:645-656.

119. Cheng H, Dufu K, Lee CS, Hsu JL, Dias A, Reed R. Human mRNA Export Machinery Recruited to the 5′ End of mRNA. Cell 2006; 127:1389-1400.

120. Nojima T, Hirose T, Kimura H, Hagiwara M. The Interaction between Cap-binding Complex and RNA Export Factor Is Required for Intronless mRNA Export. J Biol Chem 2007; 282:15645-15651.

121. Chi B, Wang K, Du Y, Gui B, Chang X, Wang L, *et al*. A Sub-Element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. Nucleic Acids Res 2014; 42:7305-7318.

122. Gebhardt A, Habjan M, Benda C, Meiler A, Haas DA, Hein MY, *et al*. mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. Nat Commun 2015; 6:8192.

123. Will CL, Lührmann R. Spliceosome Structure and Function. Cold Spring Harb Perspect Biol 2011; 3:a003707.

124. Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ. Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat Rev Mol Cell Biol 2013; 14:153-165.

125. Matera AG, Wang Z. A day in the life of the spliceosome. Nat Rev Mol Cell Biol 2014; 15:108-121.

126. Raczynska KD, Simpson CG, Ciesiolka A, Szewc L, Lewandowska D, McNicol J, *et al*. Involvement of the nuclear cap-binding protein complex in alternative splicing in Arabidopsis thaliana. Nucleic Acids Res 2010; 38:265-278.

127. Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj IW. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. Genes Dev 1996; 10:1683-1698.

128. O'Mullane L, Eperon IC. The Pre-mRNA 5′ Cap Determines Whether U6 Small Nuclear RNA Succeeds U1 Small Nuclear Ribonucleoprotein Particle at 5′ Splice Sites. Mol Cell Biol 1998; 18:7510-7520.

129. Lenasi T, Peterlin BM, Barboric M. Cap-binding Protein Complex Links Pre-mRNA Capping to Transcription Elongation and Alternative Splicing through Positive Transcription Elongation Factor b (P-TEFb). J Biol Chem 2011; 286:22758-22768.

130. Pabis M, Neufeld N, Steiner MC, Bojic T, Shav-Tal Y, Neugebauer KM. The nuclear cap-binding complex interacts with the U4/U6·U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. RNA 2013; 19:1054-1063.

131. Zhou Z, Licklider LJ, Gygi SP, Reed R. Comprehensive proteomic analysis of the human spliceosome. Nature 2002; 419:182-185.

132. Rappsilber J, Ryder U, Lamond AI, Mann M. Large-Scale Proteomic Analysis of the Human Spliceosome. Genome Res 2002; 12:1231-1245.

133. Martienssen R, Moazed D. RNAi and heterochromatin assembly. Cold Spring Harb Perspect Biol 2015; 7:a019323.

134. Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SIS, *et al*. RNAi-Mediated Targeting of Heterochromatin by the RITS Complex. Science 2004; 303:672-676.

135. Schalch T, Job G, Shanker S, Partridge JF, Joshua-Tor L. The Chp1-Tas3 core is a multifunctional platform critical for gene silencing by RITS. Nat Struct Mol Biol 2011; 18:1351-1357.

136. Jia S, Kobayashi R, Grewal SIS. Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. Nat Cell Biol 2005; 7:1007-1013.

137. Zhang K, Mosch K, Fischle W, Grewal SIS. Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol 2008; 15:381-388.

138. Ukleja M, Cuellar J, Siwaszek A, Kasprzak JM, Czarnocki-Cieciura M, Bujnicki JM, *et al*. The architecture of the Schizosaccharomyces pombe CCR4-NOT complex. Nat Commun 2016; 7:10433.

139. Cotobal C, Rodríguez-López M, Duncan C, Hasan A, Yamashita A, Yamamoto M, *et al*. Role of Ccr4-Not complex in heterochromatin formation at meiotic genes and subtelomeres in fission yeast. Epigenetics Chromatin 2015; 8:28.

140. Morgunova V, Akulenko N, Radion E, Olovnikov I, Abramov Y, Olenina LV, *et al*. Telomeric repeat silencing in germ cells is essential for early development in Drosophila. Nucleic Acids Res 2015; 43:8762-8773.

141. Czech B, Preall JB, McGinn J, Hannon GJ. A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway. Mol Cell 2013; 50:749-761.

142. Harigaya Y, Tanaka H, Yamanaka S, Tanaka K, Watanabe Y, Tsutsumi C, *et al*. Selective elimination of messenger RNA prevents an incidence of untimely meiosis. Nature 2006; 442:45-50.

143. Zofall M, Yamanaka S, Reyes-Turcu FE, Zhang K, Rubin C, Grewal SIS. RNA Elimination Machinery Targeting Meiotic mRNAs Promotes Facultative Heterochromatin Formation. Science 2012; 335:96-100.

144. Hiriart E, Vavasseur A, Touat-Todeschini L, Yamashita A, Gilquin B, Lambert E, *et al*. Mmi1 RNA surveillance machinery directs RNAi complex RITS to specific meiotic genes in fission yeast. EMBO J 2012; 31:2296-2308.

145. Scheibe M, Arnoult N, Kappei D, Buchholz F, Decottignies A, Butter F, *et al*. Quantitative interaction screen of telomeric repeat-containing RNA reveals novel TERRA regulators. Genome Res 2013; 23:2149-2157.

146. Rippe K, Luke B. TERRA and the state of the telomere. Nat Struct Mol Biol 2015; 22:853-858.

147. Porro A, Feuerhahn S, Delafontaine J, Riethman H, Rougemont J,

Lingner J. Functional characterization of the TERRA transcriptome at damaged telomeres. Nat Commun 2014; 5:5379.

148. Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM. TERRA RNA Binding to TRF2 Facilitates Heterochromatin Formation and ORC Recruitment at Telomeres. Mol Cell 2009; 35:403-413.

149. Arnoult N, Van Beneden A, Decottignies A. Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1α. Nat Struct Mol Biol 2012; 19:948-956.

150. Liu J, Jung C, Xu J, Wang H, Deng S, Bernad L, *et al*. Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. Plant Cell 2012; 24:4333-4345.

151. Fantes J, Ragge NK, Lynch SA, McGill NI, Collin JRO, Howard-Peebles PN, *et al*. Mutations in SOX2 cause anophthalmia. Nat Genet 2003; 33:462-463.

152. Shahryari A, Rafiee MR, Fouani Y, Oliae NA, Samaei NM, Shafiee M, *et al*. Two Novel Splice Variants of SOX2OT, SOX2OT-S1, and SOX2OT-S2 are Coupregulated with SOX2 and OCT4 in Esophageal Squamous Cell Carcinoma. STEM CELLS 2014; 32:126-134.

153. Askarian-Amiri ME, Seyfoddin V, Smart CE, Wang J, Kim JE, Hansji H, *et al*. Emerging Role of Long Non-Coding RNA SOX2OT in SOX2 Regulation in Breast Cancer. PLOS ONE 2014; 9:e102140.

154. Shahryari A, Jazi MS, Samaei NM, Mowla SJ. Long non-coding RNA SOX2OT: expression signature, splicing patterns, and emerging roles in pluripotency and tumorigenesis. Front Genet 2015; 6:196.

155. Calero G, Wilson KF, Ly T, Rios-Steiner JL, Clardy JC, Cerione RA. Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. Nat Struct Mol Biol 2002; 9:912-917.

156. Jayachandran U, Grey H, Cook AG. Nuclear factor 90 uses an ADAR2-like binding mode to recognize specific bases in dsRNA. Nucleic Acids Res 2016; 44:1924-1936.