

## SURVEY AND SUMMARY

# Post-transcriptional control of cellular differentiation by the RNA exosome complex

Isabela Fraga de Andrade<sup>†</sup>, Charu Mehta<sup>†</sup> and Emery H. Bresnick<sup>✉\*</sup>

Wisconsin Blood Cancer Research Institute, Department of Cell and Regenerative Biology, Carbone Cancer Center, University of Wisconsin School of Medicine and Public Health, 1111 Highland Avenue, 4009 WIMR, Madison, WI 53705, USA

Received August 27, 2020; Revised September 21, 2020; Editorial Decision September 23, 2020; Accepted September 30, 2020

### ABSTRACT

Given the complexity of intracellular RNA ensembles and vast phenotypic remodeling intrinsic to cellular differentiation, it is instructive to consider the role of RNA regulatory machinery in controlling differentiation. Dynamic post-transcriptional regulation of protein-coding and non-coding transcripts is vital for establishing and maintaining proteomes that enable or oppose differentiation. By contrast to extensively studied transcriptional mechanisms governing differentiation, many questions remain unanswered regarding the involvement of post-transcriptional mechanisms. Through its catalytic activity to selectively process or degrade RNAs, the RNA exosome complex dictates the levels of RNAs comprising multiple RNA classes, thereby regulating chromatin structure, gene expression and differentiation. Although the RNA exosome would be expected to control diverse biological processes, studies to elucidate its biological functions and how it integrates into, or functions in parallel with, cell type-specific transcriptional mechanisms are in their infancy. Mechanistic analyses have demonstrated that the RNA exosome confers expression of a differentiation regulatory receptor tyrosine kinase, downregulates the telomerase RNA component TERC, confers genomic stability and promotes DNA repair, which have considerable physiological and pathological implications. In this review, we address how a broadly operational RNA regulatory complex interfaces with cell type-specific machinery to control cellular differentiation.

### INTRODUCTION

#### Post-transcriptional control of cellular differentiation programs

Discovering paths taken by stem and progenitor cells to generate differentiated cell progeny continues to represent a productive line of investigation, and answering fundamental mechanistic questions on this problem will almost certainly spawn innovative biomedical applications. As a general principle, intrinsic and microenvironment mechanisms dynamically control cell fate decisions. With a precursor cell competent for multi- or unilineage differentiation, extracellular signaling and intracellular signaling establish regulatory networks that trigger massive phenotypic (e.g. transcriptome and proteome) remodeling as a vital component of the differentiation process. While transcriptional networks associated with stem and progenitor cell differentiation have been studied extensively (1–5), and post-transcriptional mechanisms are implicated in differentiation (6–10), how RNA regulatory complexes control differentiation by decreasing select protein-coding and non-coding RNAs (ncRNAs), while allowing others to accumulate is not thoroughly defined.

Multi-omic strategies that merge proteomic and transcriptomic datasets to discover differentiation mechanisms often lead to a focus on concordant regulation of RNA and protein. However, technical and biological parameters create considerable discordance. From a technical perspective, modern proteomic methodologies sample a proteome to yield rigorous data on ~10 000 proteins (11–13). When considering estimates of the constitution of the mammalian cell proteome, especially considering protein isoforms termed proteoforms (14,15), current technologies do not comprehensively identify proteome components. By contrast, next-generation sequencing-based RNA quantitation in cell populations is much more comprehensive, yield-

\*To whom correspondence should be addressed. Tel: +1 608 265 6446; Email: ehbresni@wisc.edu

<sup>†</sup>The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

ing many thousands of transcripts (16). Biologically, it is reasonable to assume that discordance reflects a profound contribution of post-transcriptional RNA regulatory mechanisms to proteome composition and cellular regulation. Rigorous evidence has emerged that the RNA-regulatory exosome complex (RNA exosome), a major component of post-transcriptional machinery, controls differentiation by exerting critical functions to shape transcriptomes and proteomes.

### RNA exosome structure/function

The RNA exosome, named after its 'exo'nucleolytic activity (17), and secretory vesicles termed exosomes are entirely different entities. Studies in *Saccharomyces cerevisiae* identified a critical RNA exosome function to process 5.8S ribosomal RNA (rRNA) from the precursor 7S rRNA (17,18), which is conserved from mice to humans (19–22). However, the scope of RNA exosome functions is considerably greater than rRNA processing, as the 3'–5' RNA exonucleolytic activity mediates quality control, processing and degradation of select protein-coding and non-coding transcripts. The RNA exosome processes and/or degrades transcripts generated from pervasive transcription that occurs throughout eukaryotic genomes, and such transcripts can exert biologically important activities (23–25).

RNA exosome catalytic activity is conferred by multiple catalytic subunits residing in a complex containing nine structural subunits: EXOSC1 (Csl4), EXOSC2 (Rrp4), EXOSC3 (Rrp40), EXOSC4 (Rrp41), EXOSC5 (Rrp46), EXOSC6 (Mtr3), EXOSC7 (Rrp42), EXOSC8 (Rrp43) and EXOSC9 (Rrp45) (26,27) (Figures 1 and 2). The RNA exosome protein components and complex structure are highly conserved in eukaryotes (28,29). Six subunits (EXOSC4–EXOSC9) generate a barrel-like hexameric structure that creates a scaffold for the RNA substrate. EXOSC1, EXOSC2 and EXOSC3 form a trimeric central pore that caps the barrel and has RNA-binding activity via S1 and KH domains within these proteins. This central pore recruits RNA and protein cofactors to the complex (30,31).

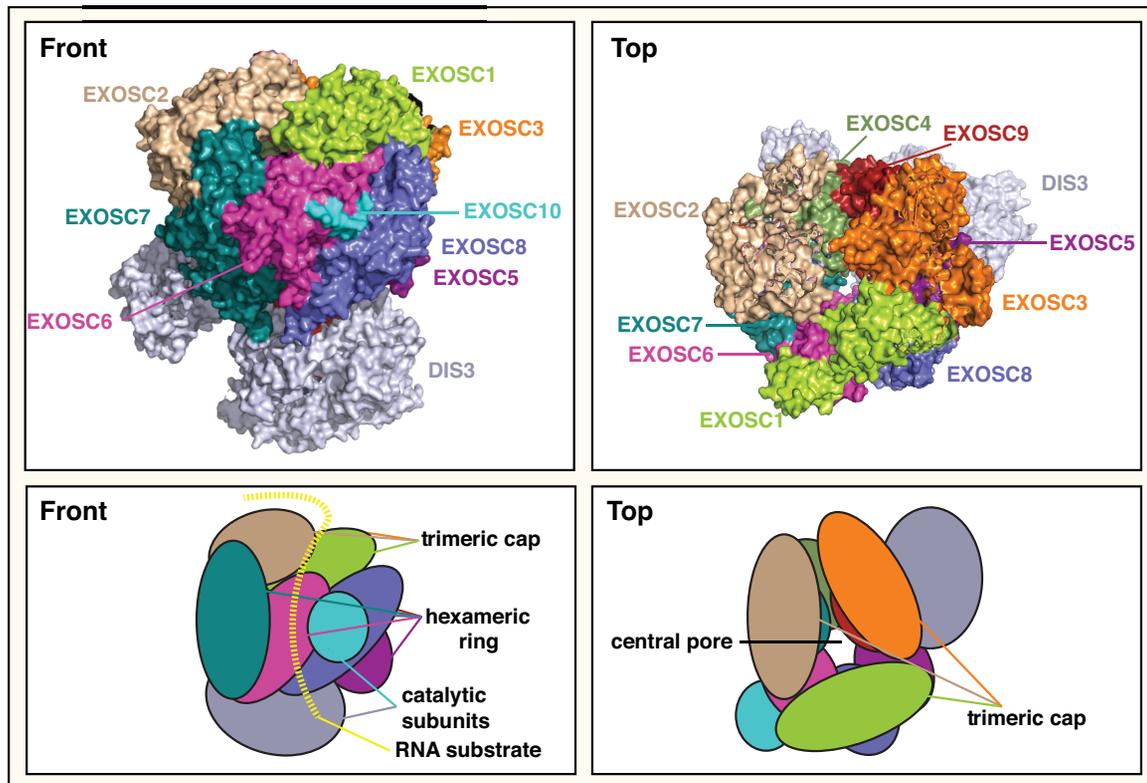
Human cells express three orthologs (DIS3, DIS3L and DIS3L2) of the yeast catalytic subunit Rrp44 that bind the inactive core (30,32–35). An additional catalytic subunit, Rrp6/EXOSC10, binds the core (19,33–36). In HEK293 and HeLa cells, differences in the subcellular localization of the catalytic subunits have been described. Subcellular fractionation and immunofluorescence analyses revealed that DIS3 is predominantly nuclear, DIS3L is cytoplasmic and EXOSC10 is enriched in the nucleolus (37–39). DIS3L2 lacks the PIN domain that mediates RNA exosome binding and is not considered to be an RNA exosome component (40). DIS3L2 mediates processing of cytoplasmic 5.8S rRNA intermediates and aberrant ncRNA species (22,41). Whereas DIS3, DIS3L and EXOSC10 all have 3'–5' exonucleolytic activity, DIS3 is the only catalytic subunit with endonucleolytic activity (38). EXOSC10 processes the 5.8S rRNA precursor within the nucleolus, and DIS3 degrades and processes protein-coding and non-coding nuclear transcripts (21,23,25).

To assess the consequences of rapidly depleting DIS3 and EXOSC10, an auxin-inducible degron was deployed in

HCT116 colon cancer cells (21). While both catalytic subunits are essential for colony-forming activity, which involves a 10-day assay, acute DIS3 depletion (1 h) has the greatest impact on the transcriptome. DIS3 depletion upregulated average reads from 4701 promoter upstream transcripts (PROMPTs), 960 enhancer RNAs (eRNAs), and 4356 premature cleavage and polyadenylation products. By contrast, EXOSC10 depletion for 1 h did not alter these transcripts. A 24- and 48-h EXOSC10 depletion modestly increased PROMPTs. Nuclear transcript regulation is often attributed to DIS3, yet its depletion triggers EXOSC10 nuclear localization, and loss of both catalytic subunits in HCT116 colon cancer cells synergistically upregulates PROMPTs (21). In certain contexts, DIS3 and EXOSC10 share functions (21), yet whether this principle can be extrapolated to multiple systems is unclear.

How can a protein complex negotiate vast RNA ensembles to exert prescriptive actions on select RNAs? RNA exosome-binding protein cofactors are important specificity determinants, and RNA exosome complex participation in diverse complexes contributes to its broad substrate targeting capacity. The RNA helicase hMTR4 allows the RNA exosome to access substrates. In yeast and human cells, MTR4 promotes RNA exosome function (42–44) and recruits additional protein cofactors that select target RNAs based on molecular attributes, including cap, poly(A) tail and transcript length (45). In HeLa and HEK293 cells, RNA exosome interactions with hMTR4, ZCCHC7 and PAPD5 form the nucleolar TRAMP complex that functions in pre-rRNA and small nucleolar RNA processing. The RNA exosome also interacts with hMTR4, RBM7 and ZCCHC8 to form the nuclear exosome targeting (NEXT) complex that regulates nuclear PROMPTs, eRNAs, unprocessed 3' extended small nuclear RNAs (snRNAs), histone RNAs, and transcripts derived from long intergenic nuclear elements and long terminal repeats (31,46–48). In HeLa cells, the nuclear RNA exosome associates with hMTR4, ZFC3H1 and the poly(A) binding protein PABPN1 to form the PAXT [poly(A) exosome targeting] connection that degrades long poly(A)-containing PROMPTs, long non-coding RNAs (lncRNAs), eRNAs and prematurely terminated transcripts (45,49). Dynamic control of subnuclear localization is likely a crucial parameter that will be accessible with high-resolution imaging strategies.

While the qualitative and quantitative attributes of RNA ensemble remodeling infer an important RNA exosome function to control cellular differentiation, the remodeling might merely be a consequence of primary regulatory mechanisms. Unlike the sophisticated knowledge of RNA exosome biochemical and molecular mechanisms that have been elegantly elucidated at the atomic level (28–30,42,44,50–59), many questions remain unanswered regarding RNA exosome biological functions and how cellular and physiological contexts impact RNA exosome structure/function. Do cell type-specific regulatory milieus create and/or require unique mechanistic permutations, or can the canonical RNA exosome paradigm be applied to any system? Is the RNA exosome essential for survival and/or function in all cell types, in which case its disruption would be universally catastrophic, or does the RNA exo-



**Figure 1.** RNA exosome structure at a resolution of 3.45 Å determined by Lima and colleagues (28) using cryo-electron microscopy. Modified from PDB ID: 6D6Q.

some exert more specialized roles in certain cell types? Are RNA exosome functions, e.g. rRNA processing, uniquely important in highly proliferative cells versus terminally differentiated or quiescent cells, or do non-proliferating cells critically rely on the RNA exosome? Since differentiation requires seamless integration of cell-extrinsic and -intrinsic mechanisms, is the RNA exosome a downstream effector of developmental signaling pathways? Our goal is to present current knowledge of RNA exosome mechanisms in diverse cellular contexts, evidence supporting its role in controlling differentiation, knowledge gaps vis-à-vis the questions posed above and insights from human genetics.

### RNA EXOSOME-DEPENDENT MECHANISMS TO CONTROL CELLULAR DIFFERENTIATION

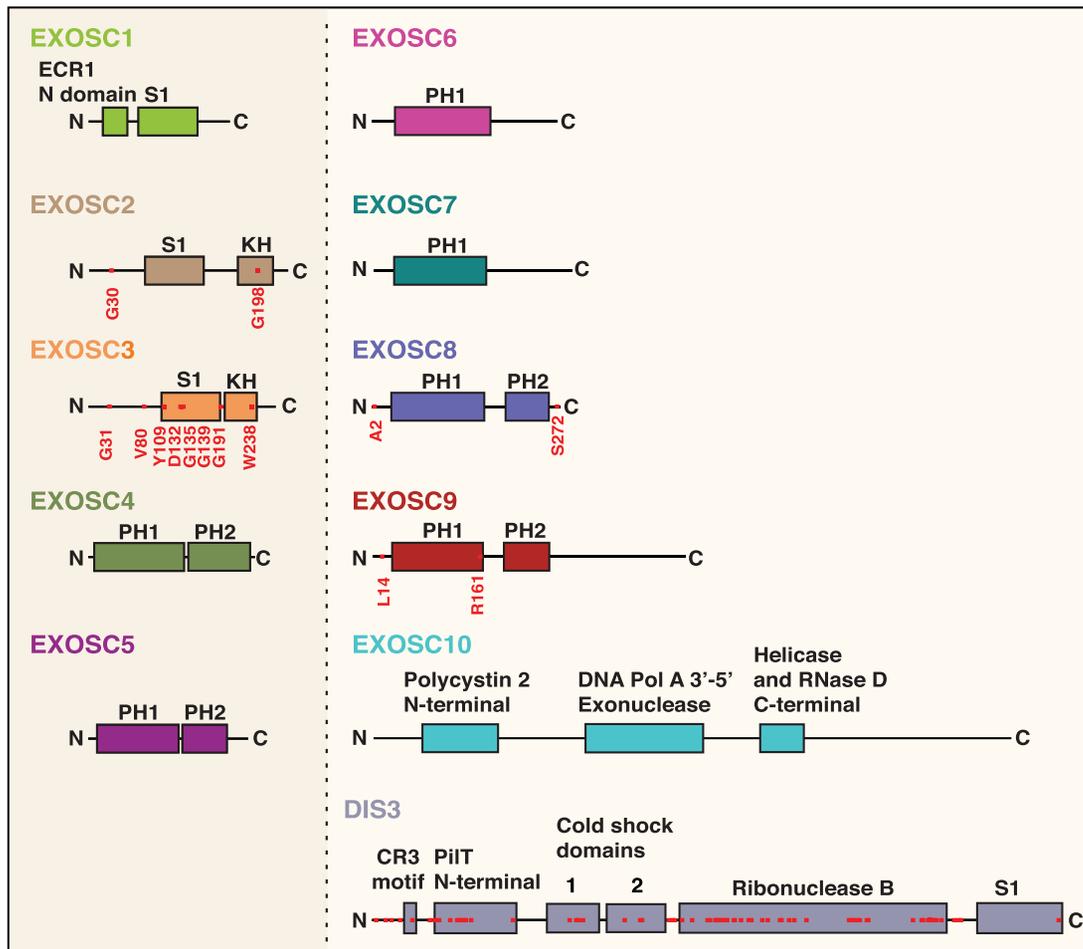
Ensuring normal RNA exosome expression and activity prevents aberrant RNA processing and accumulation of non-physiological levels of transcripts and therefore is vital for establishing and maintaining cellular functions. RNA exosome loss-of-function studies revealed upregulation of eRNAs, lncRNAs and PROMPTs, which have short half-lives due to co-transcriptional degradation. The RNA exosome eliminates these products of pervasive transcription, which is common in mammalian genomes, and RNA exosome disruption facilitates their detection. Ectopically elevated RNAs can promote genomic instability (60,61), activate antiviral defense mechanisms (62) and compete for binding of factors regulating diverse nuclear processes, including transcription (25,60,62,63). Given the

cell type specificity of gene expression mechanisms, developing principles that explain the physiological and pathological mechanisms requires analyses in diverse biological contexts.

### Skin epidermal differentiation

The scope and diversity of the RNA exosome's RNA client portfolio portend that loss-of-function perturbations of the RNA exosome would be lethal in all cell types. However, this is not the case with human skin epidermal progenitor cells. Studies on mechanisms governing skin epidermal development revealed that the RNA exosome maintains the undifferentiated state of epidermal progenitors (64,65). The RNA exosome subunits EXOSC2, EXOSC3, EXOSC7 and EXOSC9 are highly expressed in human epidermal progenitors, and their expression declines during differentiation. Depleting structural subunits EXOSC7 or EXOSC9, or the catalytic subunit EXOSC10, with shRNAs causes cell cycle arrest, impaired proliferation and premature differentiation. These phenotypes reflect a cell-autonomous RNA exosome requirement in progenitors.

EXOSC9 depletion alters expression of ~600 genes, including upregulation of differentiation-associated genes and downregulation of proliferation-linked genes (65). The loss of proliferation-linked genes correlates with a proliferation defect. EXOSC9 depletion increases expression of *GRHL3*, encoding a transcription factor necessary for epidermal differentiation and epidermal barrier formation during mouse embryogenesis (66,67). *GRHL3* is a mem-



**Figure 2.** RNA exosome subunit domain organization. Protein domains were identified from InterPro (<https://www.ebi.ac.uk/interpro/search/text/>). The relative sizes of subunits are shown, and human disease mutations are depicted as red dots (123,126–128,135,137,140,147,148). The trimeric cap proteins EXOSC1, EXOSC2 and EXOSC3 contain S1 and/or the K homology (KH) domains that mediate RNA binding. The inactive barrel proteins EXOSC4–EXOSC9 contain PH1 and/or PH2 domains that mediate protein–protein binding. EXOSC10 and DIS3 catalytic subunits contain exoribonucleolytic domains, DNA Pol A 3′–5′ exonuclease and ribonuclease B, respectively. EXOSC10 also contains a polycystin 2 N-terminal (PMC2NT) domain that interacts with CID yeast homolog (33) and a helicase and RNase D C-terminal (HRDC) domain proposed to have RNA-binding activity (34). In DIS3, the CR3 motif composed of three cysteine residues is functionally important (35), PiIT N-terminal domain (PIN) imparts its endonuclease activity (37) and cold shock domains are not functionally characterized in this context.

ber of a family of mammalian transcription factors (68) related to the *Drosophila* grainyhead transcription factor, which is critical for pattern formation and multiple aspects of *Drosophila* development (69). EXOSC9 binds and degrades *GRHL3* mRNA. As epidermal progenitors mature, RNA exosome components decrease, *GRHL3* transcripts accumulate and *GRHL3* promotes differentiation (64). EXOSC9-dependent skin development has also been described in *Xenopus* (70).

The RNA exosome-dependent mechanism to control epidermal progenitor differentiation is attributed to regulation of the levels of a single transcription factor. Considering that several hundred transcripts are sensitive to RNA exosome disruption, this mechanism may have multiple functionally important layers involving direct and indirect consequences of RNA exosome actions. These layers may have a hierarchical importance, with their integration ensuring normal differentiation. *GRHL3* regulation may be a primary step reinforced by other RNA alter-

tations that facilitate epidermal differentiation and/or suppress cell state transitions incompatible with efficient epidermal differentiation. Alternatively, the aggregate alterations in RNA exosome-regulated RNA ensembles may collectively instigate or promote a pro-differentiation cell state transition.

### Erythrocyte differentiation

A paradigm for RNA exosome-regulated progenitor cell differentiation has also emerged from studies of erythropoiesis in which hematopoietic stem cell (HSC)-derived, lineage-committed progenitors generate immature erythroblasts that progressively differentiate into enucleated erythrocytes (1). Studying this process not only reveals mechanisms of differentiation, but also provides insights into pathologies involving ‘ineffective erythropoiesis’, e.g. myelodysplastic syndrome, a pre-leukemia disease that can be refractory to therapeutic agents that promote erythro-

poiesis (71). Given the vital function of erythrocytes to protect cells and tissues from hypoxic damage, insufficient erythrocyte generation and function deleteriously impact diverse cell types and organ function (72).

One of six mammalian GATA transcription factors, GATA1, controls the differentiation of hematopoietic progenitor cells into erythrocytes (73,74). In a genetic rescue assay with GATA1-null murine erythroid precursor cells, which mimic a normal immature adult erythroid cell (proerythroblast) and express a conditional GATA1 allele (75,76), GATA1 collaborates with the forkhead transcription factor FOXO3 to repress expression of multiple RNA exosome subunit genes (77,78). Consistent with this finding, during primary mouse and human erythroid cell differentiation, RNA exosome subunit genes are transcriptionally repressed. Depleting the RNA exosome structural subunits EXOSC8 or EXOSC9 with shRNAs from primary mouse fetal liver hematopoietic progenitors reduces RNA exosome stability and disrupts the balance between erythroid precursor proliferation and differentiation to favor differentiation (78).

As in the epidermal system, RNA exosome maintains the undifferentiated state, and this activity involves sustained expression of the stem cell factor (SCF)-activated transmembrane receptor tyrosine kinase c-Kit (78). SCF-dependent c-Kit signaling regulates erythroid progenitor proliferation and survival. Depleting individual RNA exosome subunits reduces c-Kit expression and SCF-dependent c-Kit signaling, and erythroid precursors precociously acquire responsiveness to a distinct cytokine, erythropoietin (Epo), which acts through the erythropoietin transmembrane receptor (EpoR) to promote erythroblast survival and differentiation (78). GATA1 has a dual impact on c-Kit, involving direct GATA1 repression of *Kit* transcription (79) and GATA1 repression of RNA exosome subunit genes, which results in reduced c-Kit levels (78) (Figure 3). GATA1 is implicated in increasing *EpoR* gene expression (80,81). RNA exosome subunit depletion in fetal liver-derived primary erythroid precursors induces cell cycle arrest and erythroid differentiation (77). Beyond restricting differentiation, the RNA exosome supports highly proliferative erythroid progenitor cells, termed burst-forming unit-erythroid (BFU-E), based on their capacity to form diagnostic colonies in a methylcellulose-based colony assay. RNA exosome depletion quantitatively eliminates BFU-E. Thus, by controlling expression of a critical signaling molecule, c-Kit, and perhaps through additional functions, the RNA exosome sustains progenitor proliferation and opposes differentiation (Figure 3). As *Exosc3*<sup>-/-</sup> murine embryonic stem cells (mESCs) express *Kit* 5.9-fold lower than wild-type cells (25), this mechanism might operate in other contexts.

How does RNA exosome disruption abrogate *Kit* expression? In a canonical RNA degradation mechanism, RNA exosome disruption would lead to elevated levels of RNAs targeted by the RNA exosome. However, RNA exosome disruption decreases *Kit* primary transcripts and mRNA. As EXOSC9 occupies the *Kit* promoter, coding region and 3' UTR, the RNA exosome might function directly at the *Kit* locus through an unresolved mechanism, although an indirect mechanism cannot be ruled out (78).

In erythroid cells, the RNA exosome regulates mRNAs and primary transcripts of genes that establish erythrocyte functions, such as *Alas2*, encoding the rate-limiting step in heme biosynthesis, *Hbb-b1*, encoding the hemoglobin  $\beta$  chain, and *Slc4a1*, encoding an anion transporter (77). Primary transcript regulation and EXOSC9 occupancy at the respective promoters suggest that the RNA exosome either regulates transcription of these genes or exerts non-transcriptional functions directly at these loci.

The RNA exosome interacts with stalled and backtracked POL II and occupies certain active chromatin sites (82). Ongoing transcription can attract the RNA exosome to genomic loci. In B cells, EXOSC3 or EXOSC10 depletion reduces levels of chromatin components (H3K9me2 and HP1 $\gamma$ ) often linked to repression at RNA exosome-regulated eRNA loci (24,25). Whether these responses reflect direct RNA exosome actions or indirect regulation via post-transcriptional RNA exosome functions is unclear. Nevertheless, the RNA exosome regulates gene expression by processing or degrading RNAs and/or by modulating transcription—directly or indirectly.

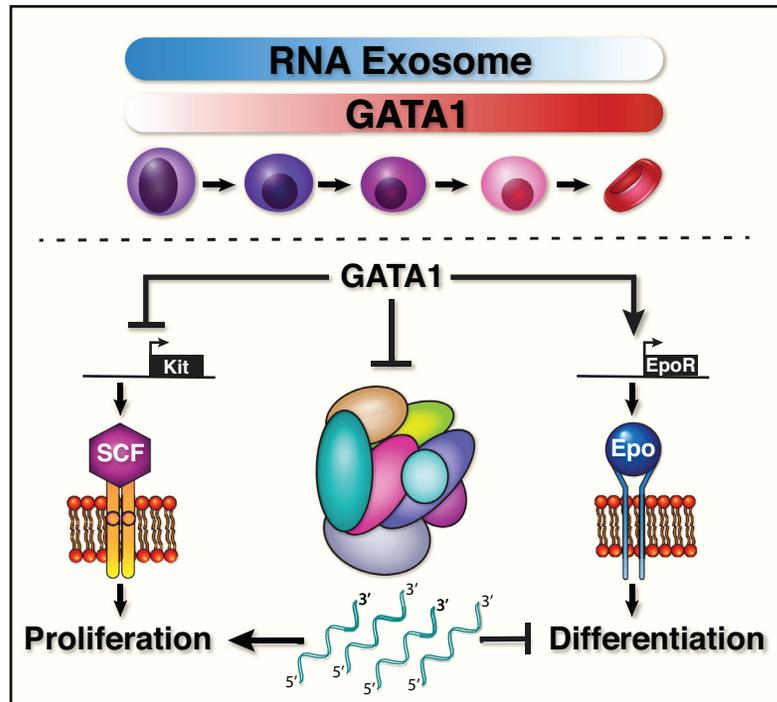
### Embryonic stem cell differentiation

Studies in induced pluripotent stem cells (iPSCs) and ESCs have revealed a process reminiscent of the GATA1–RNA exosome paradigm, in which a transcriptional mechanism is tightly coupled to a post-transcriptional mechanism to balance progenitor proliferation with differentiation. The zinc finger transcription factor ZSCAN10 occupies promoters and upregulates *Exosc1*, *Exosc2* and *Exosc5* expression in iPSCs (83). High RNA exosome levels favor the degradation of AU-rich element (ARE)-containing RNAs, including *Gpx2* mRNA encoding a protein that maintains redox potential and the DNA damage response (DDR). AREs are a common *cis*-element mediating RNA exosome-dependent RNA decay in mammalian cells (84–86). In human ESCs (hESCs), elevated RNA exosome subunits correlate with pluripotency, while subunit depletion induces differentiation. *EXOSC3*, *EXOSC4* and *EXOSC10* expression increases during reprogramming of human fibroblast cell lines (20).

In mESCs, NEXT complex disruption (48) and dismantling PAXT by depleting *Zfc3h1* (63) suppress differentiation. Polycomb repressive complex 2 (PRC2) binds chromatin and catalyzes H3K27me3 to repress genes, e.g. during ESC differentiation (87). Disruption of the PAXT connection, which the RNA exosome utilizes to target processed polyadenylated transcripts, elevates nuclear RNAs (45). In *Zfc3h1*<sup>-/-</sup> mESCs, excessive nuclear RNAs disrupt interaction between PRC2 subunits, reducing chromatin occupancy and decreasing H3K27me3 marks at target loci (63). These molecular defects compromise the repression of pluripotency genes, a key step in ESC differentiation.

### B-lymphocyte biology

Elegant studies in B lymphocytes have elucidated how the RNA exosome controls DNA mutagenesis that alters the B-cell genome to yield the plethora of receptors and antibodies that recognize pathogen-derived antigens (24,25,61,88).



**Figure 3.** RNA exosome controls erythroid differentiation by regulating the balance between erythroid precursor proliferation and differentiation. RNA exosome disruption downregulates *Kit* mRNA and protein, and therefore RNA exosome confers SCF-mediated receptor tyrosine kinase (c-Kit) signaling, which is vital to maintain the undifferentiated state of erythroid precursor cells (77,78). The repression of genes encoding RNA exosome subunits and the direct repression of *Kit* transcription constitute an important circuit within the GATA1-dependent genetic network that promotes erythroid differentiation. In addition to abrogating c-Kit expression and signaling, RNA exosome disruption is associated with precocious acquisition of Epo signaling. Epo binds the EpoR to drive erythroid differentiation. Although the full ensemble of transcripts directly regulated by the RNA exosome has not been described in this system, the RNA exosome post-transcriptional mechanism involves degradation of transcripts required for differentiation and accumulation of transcripts, e.g. encoding c-Kit, that support highly proliferative erythroid precursors termed BFU-E. RNA exosome disruption depletes BFU-E.

In mammals, V(D)J recombination occurs in the bone marrow and thymus during the early maturation of B and T cells to generate an initial repertoire of immunoglobulins (Igs) (89). As B cells develop and migrate to other lymphoid organs, additional mechanisms enhance Ig diversity and antigen-binding affinity. Somatic hypermutation (SHM) and class switch recombination (CSR) require the activity of activation-induced cytidine deaminase (AID), which deaminates cytidines at specific sites to recruit DNA repair machinery (90–92). AID expression in B cells is induced by antigen binding to the B-cell receptor.

The RNA exosome directly interacts with AID and functions as a co-factor to facilitate AID access to chromatin and deamination of template and non-template DNA strands (88). The inactive RNA exosome core containing structural and RNA-binding subunits, but lacking catalytic subunits, suffices to recruit AID and stimulate deaminase activity. In addition, the RNA exosome regulates AID mutagenic activity by preventing off-target genomic lesions (24,93). Studies with a mouse model expressing a conditional allele of *Exosc3* demonstrated that RNA exosome recruitment to AID genomic target loci is essential for SHM and CSR (24). RNA-seq analysis from *Exosc3*<sup>-/-</sup> primary B cells revealed antisense transcript accumulation at AID/RNA exosome target sites. Furthermore, genes expressing antisense transcripts correlated with loci involved in AID-dependent translocation events. These genes in-

clude *Myc*, *Pax5*, *Cd79b*, *Cd83* and *Pim1*, and their mutations occur in diffuse large B-cell lymphoma (94). The regulation of AID activity is therefore critical for B-cell development and activation and ensures that mutagenic activity is restricted to target loci. Non-coding RNA transcription and the RNA exosome are crucial to regulate AID function and prevent genomic aberrations that increase genomic instability (24,61).

Conditional B-cell-specific *Exosc3* knockout mice have revealed context-dependent RNA exosome functions (24). *CD19* promoter-driven targeted ablation of *Exosc3* in murine B cells arrests development at an immature stage. By contrast, *Exosc3* deletion driven by the B-cell-specific *AID* promoter does not compromise cellular survival nor proliferation, suggesting a stage-specific RNA exosome requirement in precursor B cells.

#### Telomerase regulation

RNA exosome regulation of telomerase activity has broad implications for physiological and pathological mechanisms of cellular differentiation. In yeast and human systems, the RNA exosome processes a poly(A)<sup>+</sup> precursor of telomerase RNA (TERC), the non-coding RNA template for telomerase reverse transcriptase activity (95,96). In hESCs, a functional analysis of cells harboring a human disease *Dkc1* mutation established an important link

between RNA exosome and telomerase (97). Dyskerin is a small nucleolar ribonucleoprotein that binds and stabilizes TERC, thereby increasing telomerase activity (98). Human disease mutations of *DKC1*, which encodes dyskerin, cause the bone marrow failure syndrome dyskeratosis congenita (99,100). In *DKC1*-mutant hESCs, TERC levels are low, and EXOSC3 depletion partially rescues TERC (97). Elevating TERC increases telomerase activity and telomere length, and reduces DNA damaging signaling.

Inhibiting the poly(A) polymerase PAPD5, which targets transcripts for RNA exosome-mediated processing and/or degradation and promotes TERC maturation, increases the capacity of hESCs to undergo hematopoietic differentiation (97,101). *DKC1* expression is high in erythroid, relative to myeloid, cells (102). During proerythroblast maturation, *DKC1* expression declines, analogous to RNA exosome loss. Does RNA exosome disruption promote telomere elongation universally or in a cell type-specific manner? Do post-transcriptional mechanisms counteract RNA exosome-mediated TERC degradation to establish a balance critical for telomere regulation? As the decline in dyskerin levels that is associated with proerythroblast maturation would be predicted to decrease TERC, and RNA exosome loss would increase TERC, further mechanistic analyses are required to address this problem.

### Genome stability and DNA repair

Since a genetic aberration in a stem cell genome will be transmitted to all of its progeny, including highly proliferative progenitor cells, ensuring stem cell genome integrity is of paramount importance. Thus, the machinery that repairs damaged DNA and confers genome integrity is vital for maintaining stem and progenitor cell phenotypes. Exposure of long-lived quiescent HSCs to DNA damage creates a risk to all blood cell progeny. DNA repair and anti-apoptotic mechanisms prevent malignant transformation of HSCs, while maintaining the quiescent HSC pool and avoiding bone marrow failure.

Dysregulation of RNA processing machinery negatively impacts genome stability (103). Given RNA exosome activity to regulate differentiation, it is instructive to consider DDR involvement in RNA exosome mechanisms. As described earlier, in an RNA exosome-dependent manner, AID exerts its mutagenic activity, and genomic alterations require DNA damage repair mechanisms, including non-homologous end joining, as part of CSR and SHM (89). Regions transcribing substrates targeted by AID and RNA exosome are sensitive to DNA double-strand breaks and genome instability (24). In B cells, select AID target genes can be translocated to the *IGH* locus due to AID-generated DNA breaks. The impaired AID activity of *Exosc3*-mutant B cells increases short ncRNAs transcribed in an antisense direction relative to transcription start sites, rendering these 'translocation hotspots' vulnerable to genetic alterations.

Divergent transcription occurs commonly at mammalian promoters and distal enhancers and generates transcripts from sense and antisense strands (104). In mESCs, EXOSC3 and EXOSC10 depletion increases divergent transcription at RNA exosome-regulated, eRNA-expressing loci, which elevates DNA double-strand breaks, as mea-

sured by  $\gamma$ H2AX and RNaseH-sensitive R-loops (25). R-loop hybrid RNA/DNA nucleotide structures, which span 100–2000 bp, involve a nascent RNA molecule and the DNA template from which it is transcribed, with displacement of the non-template DNA strand. R-loops have regulatory implications for all chromatin-based processes (105,106). RNA exosome exonucleolytic activity prevents DNA damage, at least in part, by removing transcriptional products that form detrimental hybrid structures, including R-loops (25). R-loops modulate protein complex binding, e.g. inhibiting PRC2 occupancy in mESCs (107). R-loop dismantling, induced by the specific RNaseH1 activity to degrade RNA/DNA hybrids, disrupts mESC differentiation (108).

An important link between the RNA exosome and R-loops emerged from studies on senataxin (SETX), the human homolog of yeast RNA/DNA helicase Sen1, which functions in transcription termination and R-loop resolution (109–111). Senataxin interacts with the RNA exosome, and human neurological disease *SETX* mutations disrupt this interaction (112,113). Similar to its activity in NEXT and PAXT complexes, the RNA exosome can partner with a helicase to access RNA in the degradation mechanism. In HeLa cells, excessive R-loop formation promotes interaction and colocalization of sumoylated EXOSC9 and SETX (113). Furthermore, R-loops enhance transcription by inhibiting binding and methylation of GC-rich promoters by DNA methyltransferase 1 (DNMT1) (110). A *SETX* gain-of-function mutation (L389S) from amyotrophic lateral sclerosis 4 (ALS4) patients decreases R-loops in the bone morphogenetic protein and membrane-bound activin inhibitor (*BAMBI*) promoter, facilitating methylation by DNMT1. Decreased *BAMBI* increases TGF- $\beta$  pathway activation, resembling the ALS4 phenotype. Given the potential broad impact of R-loops on genome function and the RNA exosome–R-loop connection, these mechanisms have broad implications for diverse biological processes.

The DDR is functionally linked to RNA surveillance machinery, including the RNA exosome. In U2OS osteosarcoma cells, the DDR signaling component MAPKAP kinase-2 (MK2) phosphorylates the RNA-binding proteins hnRNPA0 and PARN to post-transcriptionally regulate *Gadd45a* mRNA (114). RNA-binding motif protein 7 (RBM7), an essential component of the NEXT complex, is a substrate-recruiting protein for 3'–5' degradation by the nuclear RNA exosome (46). UV-induced DNA damage increases MK2-mediated RBM7 multi-site phosphorylation, resulting in RBM7 RNA-binding site sequestration by phospho-binding 14-3-3 proteins and accumulation of RNA exosome-regulated PROMPTs (115). 14-3-3 proteins function in DDR and many other regulatory contexts (116). In HCT116 colon cancer cells, other RNA surveillance components, ZCCHC8 and MTR4, also respond to UV-induced damage by binding to 14-3-3 proteins (115). In HeLa cells, RBM7 multi-site phosphorylation and sequestration of RNA-binding sites decreases RNA association, resulting in accumulation of PROMPTs proGADD45a, proPOGZ, proSTK11IP, proRBM39, proEXT1 and proDNAJB4, among other unknown transcripts (115,117). Thus, RBM7 is required to recruit PROMPTs for RNA exosome-mediated degradation, and DNA damage and

stress antagonize RNA exosome-mediated PROMPT degradation.

What is the relationship between RNA exosome degradation of mRNAs and MK2-mediated mRNA stabilization? In response to DNA damage stress, MK2/p38 relocalizes from the nucleus to the cytoplasm to phosphorylate RNA-binding proteins and suppress *Gadd45a* mRNA degradation (114). GADD45A promotes a positive feedback loop and maintains MK2/p38 activity to inhibit CDC25B, thus preventing cells from reentering M phase prematurely during the DDR. Several ARE-containing transcripts are stabilized in an MK2-dependent manner, including mRNAs encoding for TNF- $\alpha$  and IL-6. In primary murine erythroblasts, EXOSC8 or EXOSC9 depletion induces cell cycle arrest and expression of cell cycle regulatory genes, including ARE-containing *Gadd45a* mRNA (77). MK2 and RNA exosome might share a target mRNA cohort that functions in cell cycle regulation in response to DNA damage. *Gadd45a* encodes a member of the growth arrest and DNA damage-inducible 45 protein family that confers genome stability (118). GADD45A functions in the HSC stress response. *Gadd45a* expression is elevated by stresses, including  $\gamma$ -IR, UV-IR and cisplatin (119), and suppresses DNA damage in HSCs (120).

In mammalian cells, transcription occurring near DNA double-strand breaks can generate a class of non-coding transcripts termed damage-induced, long non-coding RNAs (dilncRNAs) (121). dilncRNAs recruit RNA processing machinery and function in the DDR. Although EXOSC10 and DIS3 are recruited to DNA damage sites in a transcription-dependent manner, only EXOSC10 is required for homologous recombination in the DDR. EXOSC10 depletion increases dilncRNAs and impairs recruitment of replication protein A (RPA), the single-stranded DNA binding protein critical for homologous recombination (122). Considering non-coding RNA diversity, the studies outlined in the prior sections are likely to constitute only a subset of the full repertoire of RNA exosome-dependent mechanisms to control differentiation.

## RNA EXOSOME MUTATIONS IN HUMAN DISEASES

Mutations in genes encoding RNA exosome structural subunits are linked to human neurodegenerative diseases. Mutations in EXOSC2 (G30 and G198), EXOSC3 (G31, V80, Y109, D132, G135, A139, G191 and W238), EXOSC8 (A2 and S272) and EXOSC9 (L14 and R161) have been described (123–127) (Table 1). *EXOSC3*, *EXOSC8* and *EXOSC9* mutations are associated with different subtypes of pontocerebellar hypoplasia (PCH) (PCH1B, OMIM #614678; PCH1C, OMIM #616081; and PCH1D, OMIM #618065, respectively), characterized by defective development of brainstem and cerebellum structures (124,125,127–135). *Drosophila melanogaster* has been used to model PCH1B, and this analysis revealed an increased requirement for EXOSC3/Rrp40 in neurons with aging (136). Homozygous (G30V) or compound heterozygous (G30V/G198D) *EXOSC2* mutations are associated with the neurological disorder SHRF (short stature, hearing loss, retinitis pigmentosa and distinctive facies syndrome, OMIM #617763) (135,137). Autosomal recessive PCH is

characterized by cerebellar atrophy, defective spinal motor development with or without defective pons (138–140). PCH patients exhibit decreased expression of the mutant, relative to the wild-type, subunit. *EXOSC3* (G31A) and *EXOSC8* (A2V) mutations may affect levels of other EC subunits, suggesting complex destabilization (127). Individuals with mutant *EXOSC2* have reduced EXOSC2, EXOSC3 and EXOSC10 levels. *EXOSC2* mutations are implicated in impaired autophagy in *D. melanogaster* (137). Analysis of G30V/G198D patient-derived lymphoblasts revealed increases in LC3-II/LC3-I ratio and autophagy receptor p62 levels, indicative of reduced autophagic flux (137).

RNA exosome mutations in PCH and PCH-like diseases involving early or progressive onset highlight its role in neuronal development and post-birth survival of spinal motor and cerebellar neurons. *EXOSC3* variants are found in 40% of PCH1 cases worldwide (127,134). The majority of other genes mutated in PCH subtypes are involved in processing and splicing of tRNA, snRNA and pre-mRNA (e.g. *TSEN54* in PCH2, PCH4 and PCH5; *RARS2* in PCH6; *TOE1* in PCH7; *CLP1* in PCH10) (141–145). Defective RNA metabolism caused by a homozygous *RBM7* (NEXT complex subunit) mutation causes spinal motor neuropathy (146). Mutations in senataxin (*SETX*), involved in R-loop resolution and RNA processing, are found in ataxia oculomotor apraxia 2 and ALS4 (109,112). Although neural development and function are apparently hypersensitive to RNA exosome disruption, the mechanisms underlying this hypersensitivity remain enigmatic and whether compensatory mechanisms suppress phenotypes in certain contexts is unclear. Whether this is applicable to RNA exosome-linked pathology or whether subtle deviations from the steady state are uniquely pathogenic in the nervous system requires further investigation.

Despite phenotypic similarities, the human mutations in RNA exosome structural subunits are not unique to a single domain. EXOSC2 residues affected by mutations reside in the N-terminal and KH domains (135,137). EXOSC3 has eight residues associated with PCH1B: five reside in the S1 domain, two in the N terminus and one in the KH domain (147). *EXOSC8* A2V and S272T mutations alter the N- and C-termini, respectively. EXOSC9-mutated residues L14 and R161 (premature stop codon; null allele) are in the N- and C-termini, respectively (127). Functional analysis and structural predictions suggest mutations impact inter-subunit interactions and/or disrupt structure required for RNA binding or passage through the complex (30,123,147,148). Loss-of-function studies in zebrafish confirm an RNA exosome requirement for brain development (127,128).

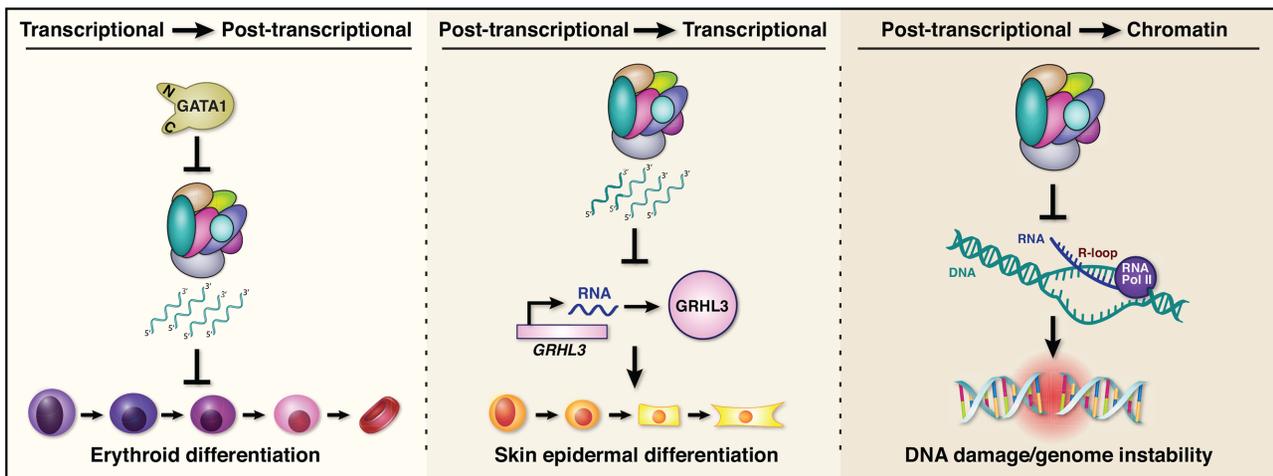
Although the RNA exosome controls rRNA biogenesis and processing (17,18,20), accumulation of rRNA intermediates was not detected in fibroblasts from a patient with a homozygous D132A mutation in *EXOSC3* and therefore it is unclear whether RNA exosome patient mutations invariably disrupt rRNA processing, or whether the disruption is context dependent (128). Hematopoiesis is particularly sensitive to ribosome perturbations (ribosomopathies) that cause bone marrow failure and anemia (149). Highly proliferative hematopoietic progenitors require a high protein biosynthetic rate, thus creating a vulnerability to even

**Table 1.** RNA exosome subunit gene mutations in human disease

RNA exosome Components	Chromosome	Protein size	Mutated Residues	Domain	aa substitutions/Genotypes	Associated disease	References		
<b>EXOSC2 (Rrp4)</b>	Chrom 9 (q34.12)	233aa (32KDa)	G30	N-terminal	G30V (hom),	SHRF	(135,137)		
			G198	KH domain	G30V/G198D (compound het)				
<b>EXOSC3 (Rrp40)</b>	Chrom 9 (p13.2)	271aa (29KDa)	G31	N-terminal	G31A (hom),	PCH1B	(128–130,134,124,125)		
			V80	N-terminal	V80F/D132A (compound het)			(132)	
			Y109	S1 domain	Y109N/D132A (compound het)			(130)	
			D132	S1 domain	D132A (hom), D132A/deletion, D132A/delins (premature STOP codon), D132A/V112I+insertion (premature STOP codon)			(128,139,130,132–134)	
			G135	S1 domain	G135E (hom)			(130)	
			A139	S1 domain	A139P/D132A (compound het)			(128)	
			G191	S1 domain	G191C (hom), G191D/D132A (compound het)			(131,140)	
			W238	KH domain	G31A/W238R (compound het)			(128,134)	
<b>EXOSC8 (Rrp43)</b>	Chrom 13 (q13.3)	276aa (30KDa)	A2	N-terminal	A2V (hom)	PCH1C	(126)		
			S272	C-terminal	S272T (hom)				
<b>EXOSC9 (Rrp45)</b>	Chrom 4 (q27)	438aa (48 KDa)	L14	N-terminal	L14P (hom)	PCH1D	(127)		
			R161	PH1 domain	L14P/R161X (compound het, X=truncated allele)				
<b>DIS3(Rrp44)</b>	Chrom13 (q21.33)	958aa (109KDa)	M1		M1I	MM	(156)		
			M1_V16del					MM	(158)
			D27		D27G			MM	(158)
			C39		C39F			MM	(162)
			L48_A56del					MM	(158)
			A56del					MM	(158)
			E81	PIN domain	E81K (2)			MM	(156,157)
			N87	PIN domain	N87K, N87S			MM	(156,158)
			T93	PIN domain	T93A			MM	(158)
			R108	PIN domain	R108C, R108S			MM	(162,158)
			H119	PIN domain	H119D			MM	(162)
			F120	PIN domain	F120L			MM	(158)
			E126	PIN domain	E126V, E126K			MM	(156)
			T131	PIN domain	T131I			MM	(152)
			G138	PIN domain	G138N			MM	(152)
			I275	Cold shock domain 1	I275R			MM	(156)
			N284	Cold shock domain 1	N284H			MM	(157)
			R285	Cold shock domain 1	R285K			MM	(162)
			D290	Cold shock domain 1	D290E			MM	(162)
			R351		R351K			MM	(162)
			T374	Cold shock domain 2	T374P (2)			MM	(153,155)
			P375	Cold shock domain 2	P375L			MM	(157)
			P412	Cold shock domain 2	P412L			MM	(152)
			R418	Cold shock domain 2	R418G			MM	(152)
			L420	Cold shock domain 2	L420V			MM	(156)
			L434*	Cold shock domain 2	nonsense			MM	(156)
			S445*		nonsense			MM	(156)
			R467	Ribonuclease B	R467P, R467Q, R467Qfs*4			MM	(156,158)
			R471	Ribonuclease B	R471W			MM	(158)
			S477	Ribonuclease B	S477R (2)			MM	(160,162)
			C483	Ribonuclease B	C483W			MM	(156)
			D485	Ribonuclease B	D485N			MM	(156)
			D487	Ribonuclease B	D487V (2), D487H			MM	(156,158)
			D488	Ribonuclease B	D488H, D488G (2), D488N (9)			MM, AML	(154,162,156–159,152)
			E501	Ribonuclease B	E501K			MM	(156)
			V504	Ribonuclease B	V504G (2)			MM	(160,162)
			F512	Ribonuclease B	F512S			MM	(157)
			R514	Ribonuclease B	R514K			AML	(159)
			A524	Ribonuclease B	A524P			MM	(156)
			G527	Ribonuclease B	G527R			MM	(156)
			Y531	Ribonuclease B	Y531D, Y531C			MM	(153,156)
			P541	Ribonuclease B	P541L			MM	(156)
			S550	Ribonuclease B	S550F, S550Y			MM	(162,158)
			N567	Ribonuclease B	N567S			MM	(158)
			H568	Ribonuclease B	H568R			MM	(158)
			K579	Ribonuclease B	K579E			MM	(158)

Table 1. Continued

RNA exosome Components	Chromosome	Protein size	Mutated Residues	Domain	aa substitutions/Genotypes	Associated disease	References
			A586	Ribonuclease B	A586V_fs*7	MM	(158)
			D604	Ribonuclease B	D604Y	MM	(157)
			M662	Ribonuclease B	M662R	MM	(162)
			E665	Ribonuclease B	E665K (2)	MM	(156,152)
			M667	Ribonuclease B	M667K, M667V	MM, AML	(159,156)
			V674	Ribonuclease B	V674L	MM	(157)
			A675	Ribonuclease B	A675T	MM	(162)
			L688	Ribonuclease B	L688R	MM	(156)
			R689	Ribonuclease B	R689Q (2)	MM	(162,158)
			A751	Ribonuclease B	A751D	MM	(158)
			H764	Ribonuclease B	H764Y (2), H764D	MM	(156,158)
			G766	Ribonuclease B	G766R (2)	MM	(160,162)
			L767	Ribonuclease B	L767F	MM	(156)
			T773	Ribonuclease B	T773I	MM	(156)
			F775	Ribonuclease B	F775L (4)	MM	(162,156–158)
			T776	Ribonuclease B	T776P	MM	(158)
			S777	Ribonuclease B	S777T, S777*	MM	(157)
			R780	Ribonuclease B	R780G, R780T (5), R780S (2), R780K (6)	MM	(160,162,152,156–158)
			R781del	Ribonuclease B		MM	(157)
			Y782	Ribonuclease B	Y782N	MM	(156)
			D784	Ribonuclease B	D784H	MM	(152)
			R789	Ribonuclease B	R789W (3)	MM	(156–158)
			H808		H808Q	MM	(156)
			C814		C814R	MM	(156)
			R820		R820W (2)	MM	(156,158)
			E929stop	S1 domain	E929stop	MM	(152)
			K952		K952T	MM	(153)



**Figure 4.** Molecular mechanisms underlying RNA exosome control of cellular differentiation and genome integrity. Three modes of RNA exosome function are depicted. Left: Transcription factors regulate expression of genes encoding RNA exosome subunits, thereby altering RNA exosome levels, which remodel transcriptomes via post-transcriptional mechanisms. The transcription factors GATA1 and ZSCAN10 repress and induce RNA exosome subunits, respectively, to regulate differentiation and maintenance of pluripotency (77,78,83). Middle: RNA exosome post-transcriptional activity regulates transcript(s) encoding a differentiation regulatory transcription factor. In human epidermal cells, the RNA exosome degrades *GRHL3* transcripts, a TF required for epidermal differentiation. In hESCs, the RNA exosome degrades transcripts that induce differentiation, and its depletion upregulates differentiation-associated factors (20,64). Right: RNA exosome exerts critical activities to maintain genome integrity. By degrading RNA molecules that form R-loop hybrid structures, as described in B cells and mESCs, the RNA exosome counteracts R-loop formation and/or maintenance, thereby regulating genome function (24,25).

modest alterations in protein synthetic capacity. However, as other cell types with high proliferative rates in the liver, gastrointestinal tract, muscle and skin are not as frequently affected, there are many unanswered questions (150). As the RNA exosome is required for erythroid progenitor activity, balancing proliferation with differentiation (77,78) and rRNA processing, presumably, its disruption impacts erythropoiesis. Since mechanisms of steady-state and stress erythropoiesis vary, in principle, phenotypes might only be manifested upon stress. Complete loss of function of any

RNA exosome subunit has not been reported in humans, and presumably, this would be lethal during embryogenesis.

Cytogenetic alterations, including translocations within the *IGH* locus, are considered to be early triggering events in multiple myeloma (MM) pathogenesis. A strong association between *DIS3* mutations and *IGH* translocations in MM patients suggests that RNA exosome-regulated AID activity may be important in this context (24,151,152). Mutations in *DIS3*, but not other RNA exosome subunit genes, occur in MM (153–159). *DIS3* somatic mutations

were detected in 4 patients in a cohort of 38 (160). Four residues were mutated within the conserved RNB exonucleolytic domain. Other studies detected *DIS3* mutations in human MM (~10%), medulloblastoma, acute myeloid leukemia and colorectal cancer (152,161–163). Although most of these mutations are somatic, germline mutations were detected in familial MM (164). Functional analyses in HEK293 cells expressing mutant *DIS3* revealed that certain disease mutations impair RNA exosome function, increasing rRNA intermediates and PROMPTs. In yeast, the mutations do not alter *DIS3* expression levels nor subunit interactions, suggesting that mutant *DIS3* assembles into the RNA exosome, but is functionally defective (161). *DIS3* mutations may inform treatment response and prognosis in MM patients (152,157).

The human genetics has great potential to unveil significant mechanistic insights that generate new RNA exosome paradigms. However, more studies are required to elucidate the molecular and cellular consequences of patient mutations and establish whether altered cellular differentiation is a functionally important component of the pathogenesis mechanisms.

## CONCLUDING REMARKS

Immense research efforts on transcriptional mechanisms and transcription-dependent networks, including collaborative contributions from the large ENCODE consortium (165,166), have unveiled a wealth of knowledge on genome function and biomedical applications. Considering the profound impact of RNA regulatory machinery on the resulting transcriptomes, a total immersion into post-transcriptional mechanisms has considerable potential to be similarly transformative. As summarized in this review, RNA exosome-regulated post-transcriptional mechanisms control fundamental biological processes, including cellular differentiation in skin, blood and ESC systems, and this is almost certainly the tip of the iceberg.

Three principles governing RNA exosome-dependent regulation of cellular differentiation have emerged (Figure 4). Cell type-specific transcriptional mechanisms function upstream of RNA exosome-dependent post-transcriptional mechanisms to repress or activate genes encoding RNA exosome subunits, thus controlling RNA exosome levels and post-transcriptional activity to remodel the transcriptome that mediates proliferation, survival and differentiation. Second, the RNA exosome-dependent post-transcriptional mechanism can function upstream of transcriptional mechanisms, e.g. to control transcripts encoding master transcriptional regulators. Finally, RNA exosome-dependent post-transcriptional mechanisms regulate chromatin, e.g. by limiting R-loop formation and promoting genome stability and DNA repair. Further research in diverse biological systems will almost certainly unveil new mechanistic permutations that extend these non-mutually exclusive mechanisms into new realms.

Regarding how the RNA exosome controls differentiation, key questions remain. (i) To what extent are RNA exosome-dependent mechanisms cell type specific and/or context dependent? While the core principles dictating fundamental RNA exosome functions may be shared in dis-

tinct cell types, mechanistic analyses in diverse systems need to be conducted to assess whether the principles can be extrapolated to any cell type, or whether a spectrum of cell type-specific mechanisms exists. (ii) The RNA exosome exerts important functions via post-transcriptional destroying and/or processing select coding RNAs and regulates R-loop formation and genome integrity, which have profound cell biological implications. Is RNA exosome regulation of coding RNAs intimately interconnected with the R-loop regulatory mechanism to control genome integrity, or do these RNA exosome functions operate predominantly in parallel mechanisms? (iii) In this review, we described paradigms in which RNA exosome-dependent regulation of coding RNAs is intrinsic to the control of cellular differentiation. Is the RNA exosome-dependent genome integrity mechanism also vital for differentiation, and is this function particularly relevant to stem and progenitor cells in which maintaining genome integrity is of vital importance? (iv) While major progress has been made in elucidating specificity determinants of RNA exosome actions on RNAs, obtaining global perspectives on how the RNA exosome establishes and maintains select RNA cohorts requires considerable additional work. (v) Given the critical roles of cell-extrinsic mechanisms in controlling cellular differentiation, is the RNA exosome a downstream effector of developmental regulatory cellular signaling mechanisms? Alternatively, does the RNA exosome control RNAs encoding secreted regulatory factors, thereby functioning as an instigator of cell-extrinsic signaling?

There are parallels between the early stage of the transcription field in which biochemical and molecular knowledge of the protein components was exceedingly sophisticated, despite major knowledge gaps vis-à-vis how the proteins and their intricate mechanisms fit into biological and pathological programs. Extending the elegant biochemical and molecular studies on RNA regulatory machines to their post-transcriptional roles in biology and pathology represents an immensely exciting phase of research with high impact for many fields.

## FUNDING

National Institutes of Health [R01 DK113186, R01 DK50107 and P30 CA014520 to E.H.B.]; Isabela Fraga de Andrade was supported in part by funding from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). The open access publication charge for this paper has been waived by Oxford University Press—*NAR* Editorial Board members are entitled to one free paper per year in recognition of their work on behalf of the journal.

*Conflict of interest statement.* None declared.

## REFERENCES

1. Katsumura, K.R., Bresnick, E.H. and Group, G.F.M. (2017) The GATA factor revolution in hematology. *Blood*, **129**, 2092–2102.
2. Wilkinson, A.C., Nakauchi, H. and Gottgens, B. (2017) Mammalian transcription factor networks: recent advances in interrogating biological complexity. *Cell Syst.*, **5**, 319–331.
3. Ton, M.N., Guibentif, C. and Gottgens, B. (2020) Single cell genomics and developmental biology: moving beyond the generation of cell type catalogues. *Curr. Opin. Genet. Dev.*, **64**, 66–71.

4. Chasman, D., Iyer, N., Fotuhi Siahpirani, A., Estevez Silva, M., Lippmann, E., McIntosh, B., Probasco, M.D., Jiang, P., Stewart, R., Thomson, J.A. *et al.* (2019) Inferring regulatory programs governing region specificity of neuroepithelial stem cells during early hindbrain and spinal cord development. *Cell Syst.*, **9**, 167–186.
5. Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G. *et al.* (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, **122**, 947–956.
6. Crittenden, S.L., Bernstein, D.S., Bachorik, J.L., Thompson, B.E., Gallegos, M., Petcherski, A.G., Moulder, G., Barstead, R., Wickens, M. and Kimble, J. (2002) A conserved RNA-binding protein controls germline stem cells in *Caenorhabditis elegans*. *Nature*, **417**, 660–663.
7. Shin, H., Haupt, K.A., Kershner, A.M., Kroll-Conner, P., Wickens, M. and Kimble, J. (2017) SYGL-1 and LST-1 link niche signaling to PUF RNA repression for stem cell maintenance in *Caenorhabditis elegans*. *PLoS Genet.*, **13**, e1007121.
8. Chua, B.A., Van Der Werf, I., Jamieson, C. and Signer, R.A.J. (2020) Post-transcriptional regulation of homeostatic, stressed, and malignant stem cells. *Cell Stem Cell*, **26**, 138–159.
9. Blanco, S., Kurovski, A., Nichols, J., Watt, F.M., Benitah, S.A. and Frye, M. (2011) The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. *PLoS Genet.*, **7**, e1002403.
10. Pereira, B., Sousa, S., Barros, R., Carreto, L., Oliveira, P., Oliveira, C., Chartier, N.T., Plateroti, M., Rouault, J.P., Freund, J.N. *et al.* (2013) CDX2 regulation by the RNA-binding protein MEX3A: impact on intestinal differentiation and stemness. *Nucleic Acids Res.*, **41**, 3986–3999.
11. Tanimura, N., Liao, R., Wilson, G.M., Dent, M.R., Cao, M., Burstyn, J.N., Hematti, P., Liu, X., Zhang, Y., Zheng, Y. *et al.* (2018) GATA/Heme Multi-omics reveals a trace metal-dependent cellular differentiation mechanism. *Dev. Cell*, **46**, 581–594.
12. Gillespie, M.A., Pali, C.G., Sanchez-Taltavull, D., Shannon, P., Longabaugh, W.J.R., Downes, D.J., Sivaraman, K., Espinoza, H.M., Hughes, J.R., Price, N.D. *et al.* (2020) Absolute quantification of transcription factors reveals principles of gene regulation in erythropoiesis. *Mol. Cell*, **78**, 960–974.
13. Johnson, K.D., Conn, D.J., Shishkova, E., Katsumura, K.R., Liu, P., Shen, S., Ranheim, E.A., Kraus, S.G., Wang, W., Calvo, K.R. *et al.* (2020) Constructing and deconstructing GATA2-regulated cell fate programs to establish developmental trajectories. *J. Exp. Med.*, **217**, e20191526.
14. Aebersold, R., Agar, J.N., Amster, I.J., Baker, M.S., Bertozzi, C.R., Boja, E.S., Costello, C.E., Cravatt, B.F., Fenselau, C., Garcia, B.A. *et al.* (2018) How many human proteoforms are there? *Nat. Chem. Biol.*, **14**, 206–214.
15. Smith, L.M. and Kelleher, N.L. (2018) Proteoforms as the next proteomics currency. *Science*, **359**, 1106–1107.
16. Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L. and Pachter, L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.*, **31**, 46–53.
17. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. and Tollervey, D. (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. *Cell*, **91**, 457–466.
18. Mitchell, P., Petfalski, E. and Tollervey, D. (1996) The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. *Genes Dev.*, **10**, 502–513.
19. Briggs, M.W., Burkard, K.T. and Butler, J.S. (1998) Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. *J. Biol. Chem.*, **273**, 13255–13263.
20. Belair, C., Sim, S., Kim, K.Y., Tanaka, Y., Park, I.H. and Wolin, S.L. (2019) The RNA exosome nuclease complex regulates human embryonic stem cell differentiation. *J. Cell Biol.*, **218**, 2564–2582.
21. Davidson, L., Francis, L., Cordiner, R.A., Eaton, J.D., Estell, C., Macias, S., Caceres, J.F. and West, S. (2019) Rapid depletion of DIS3, EXOSC10, or XRN2 reveals the immediate impact of Exoribonucleolysis on Nuclear RNA metabolism and transcriptional control. *Cell Rep.*, **26**, 2779–2791.
22. Pirouz, M., Munafò, M., Ebrahimi, A.G., Choe, J. and Gregory, R.I. (2019) Exonuclease requirements for mammalian ribosomal RNA biogenesis and surveillance. *Nat. Struct. Mol. Biol.*, **26**, 490–500.
23. Szczepinska, T., Kalisiak, K., Tomecki, R., Labno, A., Borowski, L.S., Kulinski, T.M., Adamska, D., Kosinska, J. and Dziembowski, A. (2015) DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts. *Genome Res.*, **25**, 1622–1633.
24. Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A.N. and Basu, U. (2014) Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature*, **514**, 389–393.
25. Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation, A., Chao, J., Elliott, O., Liu, Z.P. *et al.* (2015) RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell*, **161**, 774–789.
26. Schneider, C. and Tollervey, D. (2013) Threading the barrel of the RNA exosome. *Trends Biochem. Sci.*, **38**, 485–493.
27. Lingaraju, M., Schuller, J.M., Falk, S., Gerlach, P., Bonneau, F., Basquin, J., Benda, C. and Conti, E. (2019) To Process or to Decay: A Mechanistic View of the Nuclear RNA Exosome. *Cold Spring Harb. Symp. Quant. Biol.*, **84**, 155–163.
28. Weick, E.M., Puno, M.R., Januszzyk, K., Zinder, J.C., DiMattia, M.A. and Lima, C.D. (2018) Helicase-Dependent RNA decay Illuminated by a Cryo-EM structure of a human Nuclear RNA Exosome-MTR4 complex. *Cell*, **173**, 1663–1677.
29. Schuller, J.M., Falk, S., Fromm, L., Hurt, E. and Conti, E. (2018) Structure of the nuclear exosome captured on a maturing preribosome. *Science*, **360**, 219–222.
30. Liu, Q., Greimann, J.C. and Lima, C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223–1237.
31. Kilchert, C., Wittmann, S. and Vasiljeva, L. (2016) The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.*, **17**, 227–239.
32. Dziembowski, A., Lorentzen, E., Conti, E. and Seraphin, B. (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.*, **14**, 15–22.
33. Stead, J.A., Costello, J.L., Livingstone, M.J. and Mitchell, P. (2007) The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein. *Nucleic Acids Res.*, **35**, 5556–5567.
34. Liu, Z., Macias, M.J., Bottomley, M.J., Stier, G., Linge, J.P., Nilges, M., Bork, P. and Sattler, M. (1999) The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins. *Structure*, **7**, 1557–1566.
35. Schaeffer, D., Reis, F.P., Johnson, S.J., Arraiano, C.M. and van Hoof, A. (2012) The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function. *Nucleic Acids Res.*, **40**, 9298–9307.
36. Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D. and Mitchell, P. (1999) The yeast exosome and human PM-Scl are related complexes of 3'→5' exonucleases. *Genes Dev.*, **13**, 2148–2158.
37. Staals, R.H., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J., Rajmakers, R. and Pruijn, G.J. (2010) Dis3-like 1: a novel exoribonuclease associated with the human exosome. *EMBO J.*, **29**, 2358–2367.
38. Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Draskowska, K., Pastula, A., Andersen, J.S., Stepień, P.P. *et al.* (2010) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.*, **29**, 2342–2357.
39. Schilders, G., van Dijk, E. and Pruijn, G.J. (2007) C1D and hMtr4p associate with the human exosome subunit PM/Scl-100 and are involved in pre-rRNA processing. *Nucleic Acids Res.*, **35**, 2564–2572.
40. Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H. and Dziembowski, A. (2013) Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. *EMBO J.*, **32**, 1855–1868.
41. Pirouz, M., Du, P., Munafò, M. and Gregory, R.I. (2016) Dis3l2-mediated decay is a quality control pathway for noncoding RNAs. *Cell Rep.*, **16**, 1861–1873.
42. Zinder, J.C. and Lima, C.D. (2017) Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes Dev.*, **31**, 88–100.

43. Fan, J., Kuai, B., Wu, G., Wu, X., Chi, B., Wang, L., Wang, K., Shi, Z., Zhang, H., Chen, S. *et al.* (2017) Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *EMBO J.*, **36**, 2870–2886.
44. Lingaraju, M., Johnsen, D., Schlundt, A., Langer, L.M., Basquin, J., Sattler, M., Heick Jensen, T., Falk, S. and Conti, E. (2019) The MTR4 helicase recruits nuclear adaptors of the human RNA exosome using distinct arch-interacting motifs. *Nat. Commun.*, **10**, 3393.
45. Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A. *et al.* (2016) Identification of a nuclear exosome decay pathway for processed transcripts. *Mol. Cell.*, **64**, 520–533.
46. Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A. and Jensen, T.H. (2011) Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell.*, **43**, 624–637.
47. Lubas, M., Andersen, P.R., Schein, A., Dziembowski, A., Kudla, G. and Jensen, T.H. (2015) The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. *Cell Rep.*, **10**, 178–192.
48. Wu, Y., Liu, W., Chen, J., Liu, S., Wang, M., Yang, L., Chen, C., Qi, M., Xu, Y., Qiao, Z. *et al.* (2019) Nuclear exosome targeting complex core factor Zcchc8 regulates the degradation of LINE1 RNA in early embryos and embryonic stem cells. *Cell Rep.*, **29**, 2461–2472.
49. Silla, T., Schmid, M., Dou, Y., Garland, W., Milek, M., Imami, K., Johnsen, D., Polak, P., Andersen, J.S., Selbach, M. *et al.* (2020) The human ZC3H3 and RBM26/27 proteins are critical for PAXT-mediated nuclear RNA decay. *Nucleic Acids Res.*, **48**, 2518–2530.
50. Wasmuth, E.V., Januszky, K. and Lima, C.D. (2014) Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. *Nature*, **511**, 435–439.
51. Gerlach, P., Schuller, J.M., Bonneau, F., Basquin, J., Reichelt, P., Falk, S. and Conti, E. (2018) Distinct and evolutionary conserved structural features of the human nuclear exosome complex. *Elife*, **7**, e38686.
52. Makino, D.L., Schuch, B., Stegmann, E., Baumgartner, M., Basquin, C. and Conti, E. (2015) RNA degradation paths in a 12-subunit nuclear exosome complex. *Nature*, **524**, 54–58.
53. Fromm, L., Falk, S., Flemming, D., Schuller, J.M., Thoms, M., Conti, E. and Hurt, E. (2017) Reconstitution of the complete pathway of ITS2 processing at the pre-ribosome. *Nat. Commun.*, **8**, 1787.
54. Schuch, B., Feigenbutz, M., Makino, D.L., Falk, S., Basquin, C., Mitchell, P. and Conti, E. (2014) The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. *EMBO J.*, **33**, 2829–2846.
55. Makino, D.L., Baumgartner, M. and Conti, E. (2013) Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature*, **495**, 70–75.
56. Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E. and Conti, E. (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell*, **139**, 547–559.
57. Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A. and Conti, E. (2008) Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol. Cell.*, **29**, 717–728.
58. Oddone, A., Lorentzen, E., Basquin, J., Gasch, A., Rybin, V., Conti, E. and Sattler, M. (2007) Structural and biochemical characterization of the yeast exosome component Rrp40. *EMBO Rep.*, **8**, 63–69.
59. Lorentzen, E. and Conti, E. (2005) Structural basis of 3' end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol. Cell.*, **20**, 473–481.
60. Wahba, L., Gore, S.K. and Koshland, D. (2013) The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. *Elife*, **2**, e00505.
61. Pefanis, E. and Basu, U. (2015) RNA Exosome Regulates AID DNA Mutator Activity in the B Cell Genome. *Adv. Immunol.*, **127**, 257–308.
62. Eckard, S.C., Rice, G.I., Fabre, A., Badens, C., Gray, E.E., Hartley, J.L., Crow, Y.J. and Stetson, D.B. (2014) The SKIV2L RNA exosome limits activation of the RIG-I-like receptors. *Nat. Immunol.*, **15**, 839–845.
63. Garland, W., Comet, I., Wu, M., Radziszewska, A., Rib, L., Vitting-Seerup, K., Lloret-Llinares, M., Sandelin, A., Helin, K. and Jensen, T.H. (2019) A functional link between nuclear RNA decay and transcriptional control mediated by the polycomb repressive complex 2. *Cell Rep.*, **29**, 1800–1811.
64. Mistry, D.S., Chen, Y. and Sen, G.L. (2012) Progenitor function in self-renewing human epidermis is maintained by the exosome. *Cell Stem Cell*, **11**, 127–135.
65. Li, J. and Sen, G.L. (2016) Post-transcriptional mechanisms regulating epidermal stem and progenitor cell self-renewal and differentiation. *J. Invest. Dermatol.*, **136**, 746–752.
66. Ting, S.B., Wilanowski, T., Auden, A., Hall, M., Voss, A.K., Thomas, T., Parekh, V., Cunningham, J.M. and Jane, S.M. (2003) Inositol- and folate-resistant neural tube defects in mice lacking the epithelial-specific factor Grhl-3. *Nat. Med.*, **9**, 1513–1519.
67. Caddy, J., Wilanowski, T., Darido, C., Dworkin, S., Ting, S.B., Zhao, Q., Rank, G., Auden, A., Srivastava, S., Papenfuss, T.A. *et al.* (2010) Epidermal wound repair is regulated by the planar cell polarity signaling pathway. *Dev. Cell*, **19**, 138–147.
68. Wilanowski, T., Tuckfield, A., Cerutti, L., O'Connell, S., Saint, R., Parekh, V., Tao, J., Cunningham, J.M. and Jane, S.M. (2002) A highly conserved novel family of mammalian developmental transcription factors related to Drosophila grainyhead. *Mech. Dev.*, **114**, 37–50.
69. Uy, A.E., Harrison, E.J. and Bray, S.J. (1997) Tissue-specific splicing and functions of the Drosophila transcription factor Grainyhead. *Mol. Cell Biol.*, **17**, 6727–6735.
70. Noiret, M., Mottier, S., Angrand, G., Gautier-Courteille, C., Lerivray, H., Viet, J., Paillard, L., Mereau, A., Hardy, S. and Audic, Y. (2016) Ptbp1 and Exosc9 knockdowns trigger skin stability defects through different pathways. *Dev. Biol.*, **409**, 489–501.
71. Sankaran, V.G. and Weiss, M.J. (2015) Anemia: progress in molecular mechanisms and therapies. *Nat. Med.*, **21**, 221–230.
72. Bresnick, E.H., Hewitt, K.J., Mehta, C., Keles, S., Paulson, R.F. and Johnson, K.D. (2018) Mechanisms of erythrocyte development and regeneration: implications for regeneration medicine. *Development*, **145**, dev151423.
73. Tsai, S.F., Martin, D.I., Zon, L.I., D'Andrea, A.D., Wong, G.G. and Orkin, S.H. (1989) Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature*, **339**, 446–451.
74. Evans, T. and Felsenfeld, G. (1989) The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell*, **58**, 877–885.
75. Gregory, T., Yu, C., Ma, A., Orkin, S.H., Blobel, G.A. and Weiss, M.J. (1999) GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xl expression. *Blood*, **94**, 87–96.
76. Welch, J.J., Watts, J.A., Vakoc, C.R., Yao, Y., Wang, H., Hardison, R.C., Blobel, G.A., Chodosh, L.A. and Weiss, M.J. (2004) Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood*, **104**, 3136–3147.
77. McIver, S.C., Kang, Y.A., DeVilbiss, A.W., O'Driscoll, C.A., Ouellette, J.N., Pope, N.J., Camprecios, G., Chang, C.J., Yang, D., Bouhassira, E.E. *et al.* (2014) The exosome complex establishes a barricade to erythroid maturation. *Blood*, **124**, 2285–2297.
78. McIver, S.C., Katsumura, K.R., Davids, E., Liu, P., Kang, Y.A., Yang, D. and Bresnick, E.H. (2016) Exosome complex orchestrates developmental signaling to balance proliferation and differentiation during erythropoiesis. *Elife*, **5**, e17877.
79. Munugalavadla, V., Dore, L.C., Tan, B.L., Hong, L., Vishnu, M., Weiss, M.J. and Kapur, R. (2005) Repression of c-kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. *Mol. Cell Biol.*, **25**, 6747–6759.
80. Chiba, T., Ikawa, Y. and Todokoro, K. (1991) GATA-1 transactivates erythropoietin receptor gene, and erythropoietin receptor-mediated signals enhance GATA-1 gene expression. *Nucleic Acids Res.*, **19**, 3843–3848.
81. Zon, L.I., Yousoufian, H., Mather, C., Lodish, H.F. and Orkin, S.H. (1991) Activation of the erythropoietin receptor promoter by transcription factor GATA-1. *Proc. Natl. Acad. Sci. U. S. A.*, **88**, 10638–10641.
82. Lemay, J.F., Laroche, M., Marguerat, S., Atkinson, S., Bahler, J. and Bachand, F. (2014) The RNA exosome promotes transcription termination of backtracked RNA polymerase II. *Nat. Struct. Mol. Biol.*, **21**, 919–926.
83. Skamagki, M., Zhang, C., Ross, C.A., Ananthanarayanan, A., Liu, Z., Mu, Q., Basu, U., Wang, J., Zhao, R., Li, H. *et al.* (2017) RNA

- exosome complex-mediated control of redox status in pluripotent stem cells. *Stem Cell Reports*, **9**, 1053–1061.
84. Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J., Stoeklin, G., Moroni, C., Mann, M. and Karin, M. (2001) AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell*, **107**, 451–464.
  85. Tran, H., Schilling, M., Wirbelauer, C., Hess, D. and Nagamine, Y. (2004) Facilitation of mRNA deadenylation and decay by the exosome-bound, DEXH protein RHAU. *Mol. Cell*, **13**, 101–111.
  86. Chen, C.Y. and Shyu, A.B. (1995) AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem. Sci.*, **20**, 465–470.
  87. van Mierlo, G., Veenstra, G.J.C., Vermeulen, M. and Marks, H. (2019) The Complexity of PRC2 Subcomplexes. *Trends Cell Biol.*, **29**, 660–671.
  88. Basu, U., Meng, F.L., Keim, C., Grinstein, V., Pefanis, E., Eccleston, J., Zhang, T., Myers, D., Wasserman, C.R., Wesemann, D.R. *et al.* (2011) The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell*, **144**, 353–363.
  89. Deriano, L. and Roth, D.B. (2013) Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.*, **47**, 433–455.
  90. Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O. and Honjo, T. (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.*, **274**, 18470–18476.
  91. Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. and Honjo, T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, **102**, 553–563.
  92. Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E. and Alt, F.W. (2003) Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature*, **422**, 726–730.
  93. Meng, F.L., Du, Z., Federation, A., Hu, J., Wang, Q., Kieffer-Kwon, K.R., Meyers, R.M., Amor, C., Wasserman, C.R., Neuberg, D. *et al.* (2014) Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell*, **159**, 1538–1548.
  94. Pasqualucci, L. and Dalla-Favera, R. (2018) Genetics of diffuse large B-cell lymphoma. *Blood*, **131**, 2307–2319.
  95. Coy, S., Volanakis, A., Shah, S. and Vasiljeva, L. (2013) The Sm complex is required for the processing of non-coding RNAs by the exosome. *PLoS One*, **8**, e65606.
  96. Tseng, C.K., Wang, H.F., Burns, A.M., Schroeder, M.R., Gaspari, M. and Baumann, P. (2015) Human telomerase RNA processing and quality control. *Cell Rep.*, **13**, 2232–2243.
  97. Fok, W.C., Shukla, S., Vessoni, A.T., Brenner, K.A., Parker, R., Sturgeon, C.M. and Batista, L.F.Z. (2019) Posttranscriptional modulation of TERC by PAPP5 inhibition rescues hematopoietic development in dyskeratosis congenita. *Blood*, **133**, 1308–1312.
  98. Mitchell, J.R., Wood, E. and Collins, K. (1999) A telomerase component is defective in the human disease dyskeratosis congenita. *Nature*, **402**, 551–555.
  99. Heiss, N.S., Knight, S.W., Vulliamy, T.J., Klauck, S.M., Wiemann, S., Mason, P.J., Poustka, A. and Dokal, I. (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat. Genet.*, **19**, 32–38.
  100. Batista, L.F., Pech, M.F., Zhong, F.L., Nguyen, H.N., Xie, K.T., Zaug, A.J., Crary, S.M., Choi, J., Sebastiano, V., Cherry, A. *et al.* (2011) Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. *Nature*, **474**, 399–402.
  101. Shukla, S., Jeong, H.C., Sturgeon, C.M., Parker, R. and Batista, L.F.Z. (2020) Chemical inhibition of PAPP5/7 rescues telomerase function and hematopoiesis in dyskeratosis congenita. *Blood Adv*, **4**, 2717–2722.
  102. Richards, L.A., Kumari, A., Knezevic, K., Thoms, J.A., von Jonquieres, G., Napier, C.E., Ali, Z., O'Brien, R., Marks-Bluth, J., Maritz, M.F. *et al.* (2020) DKC1 is a transcriptional target of GATA1 and drives upregulation of telomerase activity in normal human erythroblasts. *Haematologica*, **105**, 1517–1526.
  103. Chan, Y.A., Hieter, P. and Stirling, P.C. (2014) Mechanisms of genome instability induced by RNA-processing defects. *Trends Genet.*, **30**, 245–253.
  104. Wu, X. and Sharp, P.A. (2013) Divergent transcription: a driving force for new gene origination? *Cell*, **155**, 990–996.
  105. Hegazy, Y.A., Fernando, C.M. and Tran, E.J. (2020) The balancing act of R-loop biology: The good, the bad, and the ugly. *J. Biol. Chem.*, **295**, 905–913.
  106. Crossley, M.P., Bocek, M. and Cimprich, K.A. (2019) R-Loops as cellular regulators and genomic threats. *Mol. Cell*, **73**, 398–411.
  107. Skourti-Stathaki, K., Torlai Triglia, E., Warburton, M., Voigt, P., Bird, A. and Pombo, A. (2019) R-Loops enhance polycomb repression at a subset of developmental regulator genes. *Mol. Cell*, **73**, 930–945.
  108. Chen, P.B., Chen, H.V., Acharya, D., Rando, O.J. and Fazio, T.G. (2015) R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat. Struct. Mol. Biol.*, **22**, 999–1007.
  109. Skourti-Stathaki, K., Proudfoot, N.J. and Gromak, N. (2011) Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell*, **42**, 794–805.
  110. Grunseich, C., Wang, I.X., Watts, J.A., Burdick, J.T., Guber, R.D., Zhu, Z., Bruzel, A., Lanman, T., Chen, K., Schindler, A.B. *et al.* (2018) Senataxin mutation reveals how R-Loops promote transcription by blocking DNA methylation at gene promoters. *Mol. Cell*, **69**, 426–437.
  111. Abraham, K.J., Khosraviani, N., Chan, J.N.Y., Gorthi, A., Samman, A., Zhao, D.Y., Wang, M., Bokros, M., Vidya, E., Ostrowski, L.A. *et al.* (2020) Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature*, **585**, 298–302.
  112. Moreira, M.C., Klur, S., Watanabe, M., Nemeth, A.H., Le Ber, I., Moniz, J.C., Tranchant, C., Aubourg, P., Tazir, M., Schols, L. *et al.* (2004) Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat. Genet.*, **36**, 225–227.
  113. Richard, P., Feng, S. and Manley, J.L. (2013) A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes Dev.*, **27**, 2227–2232.
  114. Reinhardt, H.C., Hasskamp, P., Schmedding, I., Morandell, S., van Vugt, M.A., Wang, X., Linding, R., Ong, S.E., Weaver, D., Carr, S.A. *et al.* (2010) DNA damage activates a spatially distinct late cytoplasmic cell-cycle checkpoint network controlled by MK2-mediated RNA stabilization. *Mol. Cell*, **40**, 34–49.
  115. Blasius, M., Wagner, S.A., Choudhary, C., Bartek, J. and Jackson, S.P. (2014) A quantitative 14-3-3 interaction screen connects the nuclear exosome targeting complex to the DNA damage response. *Genes Dev.*, **28**, 1977–1982.
  116. Reinhardt, H.C. and Yaffe, M.B. (2013) Phospho-Ser/Thr-binding domains: navigating the cell cycle and DNA damage response. *Nat. Rev. Mol. Cell Biol.*, **14**, 563–580.
  117. Tiedje, C., Lubas, M., Tehrani, M., Menon, M.B., Ronkina, N., Rousseau, S., Cohen, P., Kotlyarov, A. and Gaestel, M. (2015) p38MAPK/MK2-mediated phosphorylation of RBM7 regulates the human nuclear exosome targeting complex. *RNA*, **21**, 262–278.
  118. Hollander, M.C., Sheikh, M.S., Bulavin, D.V., Lundgren, K., Augeri-Henmueller, L., Shehee, R., Molinaro, T.A., Kim, K.E., Tolosa, E., Ashwell, J.D. *et al.* (1999) Genomic instability in Gadd45a-deficient mice. *Nat. Genet.*, **23**, 176–184.
  119. Gupta, M., Gupta, S.K., Balliet, A.G., Hollander, M.C., Fornace, A.J., Hoffman, B. and Liebermann, D.A. (2005) Hematopoietic cells from Gadd45a- and Gadd45b-deficient mice are sensitized to genotoxic-stress-induced apoptosis. *Oncogene*, **24**, 7170–7179.
  120. Chen, Y., Ma, X., Zhang, M., Wang, X., Wang, C., Wang, H., Guo, P., Yuan, W., Rudolph, K.L., Zhan, Q. *et al.* (2014) Gadd45a regulates hematopoietic stem cell stress responses in mice. *Blood*, **123**, 851–862.
  121. Michelini, F., Pitchiaya, S., Vitelli, V., Sharma, S., Gioia, U., Pessina, F., Cabrini, M., Wang, Y., Capozzo, I., Iannelli, F. *et al.* (2017) Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. *Nat. Cell Biol.*, **19**, 1400–1411.
  122. Domingo-Prim, J., Endara-Coll, M., Bonath, F., Jimeno, S., Prados-Carvajal, R., Friedlander, M.R., Huertas, P. and Visa, N. (2019) EXOSC10 is required for RPA assembly and controlled DNA end resection at DNA double-strand breaks. *Nat. Commun.*, **10**, 2135.

123. Fasken, M.B., Morton, D.J., Kuiper, E.G., Jones, S.K., Leung, S.W. and Corbett, A.H. (2020) The RNA exosome and human disease. *Methods Mol. Biol.*, **2062**, 3–33.
124. Kvarnung, M., Taylan, F., Nilsson, D., Anderlid, B.M., Malmgren, H., Lagerstedt-Robinson, K., Holmberg, E., Burstedt, M., Nordenskjöld, M., Nordgren, A. *et al.* (2018) Genomic screening in rare disorders: New mutations and phenotypes, highlighting ALG14 as a novel cause of severe intellectual disability. *Clin. Genet.*, **94**, 528–537.
125. Di Giovambattista, A.P., Jácome Querejeta, I., Ventura Faci, P., Rodríguez Martínez, G. and Ramos Fuentes, F. (2017) [Familial EXOSC3-related pontocerebellar hypoplasia]. *An Pediatr (Barc)*, **86**, 284–286.
126. Boczonadi, V., Muller, J.S., Pyle, A., Munkley, J., Dor, T., Quartararo, J., Ferrero, I., Karcagi, V., Giunta, M., Polvikoski, T. *et al.* (2014) EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia. *Nat. Commun.*, **5**, 4287.
127. Burns, D.T., Donkervoort, S., Muller, J.S., Knierim, E., Bharucha-Goebel, D., Faqeih, E.A., Bell, S.K., AlFaifi, A.Y., Monies, D., Millan, F. *et al.* (2018) Variants in EXOSC9 Disrupt the RNA Exosome and Result in Cerebellar Atrophy with Spinal Motor Neuronopathy. *Am. J. Hum. Genet.*, **102**, 858–873.
128. Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schoneborn, S., Menezes, M.P., Hong, J.E., Leong, D.W., Senderek, J., Salman, M.S., Chitayat, D. *et al.* (2012) Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat. Genet.*, **44**, 704–708.
129. Schwabova, J., Brozkova, D.S., Petrak, B., Mojzisova, M., Pavlickova, K., Haberlova, J., Mrazkova, L., Hedvicakova, P., Hornofova, L., Kaluzova, M. *et al.* (2013) Homozygous EXOSC3 mutation c.92G→C, p.G31A is a founder mutation causing severe pontocerebellar hypoplasia type I among the Czech Roma. *J. Neurogenet.*, **27**, 163–169.
130. Eggens, V.R., Barth, P.G., Niermeijer, J.M., Berg, J.N., Darin, N., Dixit, A., Fluss, J., Foulds, N., Fowler, D., Hortobagyi, T. *et al.* (2014) EXOSC3 mutations in pontocerebellar hypoplasia type I: novel mutations and genotype-phenotype correlations. *Orphanet. J. Rare Dis.*, **9**, 23.
131. Halevy, A., Lerer, I., Cohen, R., Kornreich, L., Shuper, A., Gamliel, M., Zimmerman, B.E., Korabi, I., Meiner, V., Straussberg, R. *et al.* (2014) Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia. *J. Neurol.*, **261**, 2165–2169.
132. Zanni, G., Scotton, C., Passarelli, C., Fang, M., Barresi, S., Dallapiccola, B., Wu, B., Gualandi, F., Ferlini, A., Bertini, E. *et al.* (2013) Exome sequencing in a family with intellectual disability, early onset spasticity, and cerebellar atrophy detects a novel mutation in EXOSC3. *Neurogenetics*, **14**, 247–250.
133. Schottmann, G., Picker-Minh, S., Schwarz, J.M., Gill, E., Rodenburg, R.J.T., Stenzel, W., Kaindl, A.M. and Schuelke, M. (2017) Recessive mutation in EXOSC3 associates with mitochondrial dysfunction and pontocerebellar hypoplasia. *Mitochondrion*, **37**, 46–54.
134. Rudnik-Schoneborn, S., Senderek, J., Jen, J.C., Houge, G., Seeman, P., Puchmajerova, A., Graul-Neumann, L., Seidel, U., Korinthenberg, R., Kirschner, J. *et al.* (2013) Pontocerebellar hypoplasia type I: clinical spectrum and relevance of EXOSC3 mutations. *Neurology*, **80**, 438–446.
135. Di Donato, N., Neuhann, T., Kahlert, A.K., Klink, B., Hackmann, K., Neuhann, I., Novotna, B., Schallner, J., Krause, C., Glass, I.A. *et al.* (2016) Mutations in EXOSC2 are associated with a novel syndrome characterised by retinitis pigmentosa, progressive hearing loss, premature ageing, short stature, mild intellectual disability and distinctive gestalt. *J. Med. Genet.*, **53**, 419–425.
136. Morton, D.J., Jalloh, B., Kim, L., Kremisky, I., Nair, R.J., Nguyen, K.B., Rounds, J.C., Sterrett, M.C., Brown, B., Le, T. *et al.* (2020) A Drosophila model of Pontocerebellar Hypoplasia reveals a critical role for the RNA exosome in neurons. *PLoS Genet.*, **16**, e1008901.
137. Yang, X., Bayat, V., DiDonato, N., Zhao, Y., Zarnegar, B., Siprashvili, Z., Lopez-Pajares, V., Sun, T., Tao, S., Li, C. *et al.* (2020) Genetic and genomic studies of pathogenic EXOSC2 mutations in the newly described disease SHRF implicate the autophagy pathway in disease pathogenesis. *Hum. Mol. Genet.*, **29**, 541–553.
138. Rudnik-Schoneborn, S., Barth, P.G. and Zerres, K. (2014) Pontocerebellar hypoplasia. *Am. J. Med. Genet. C Semin. Med. Genet.*, **166**, 173–183.
139. Biancheri, R., Cassandrini, D., Pinto, F., Trovato, R., Di Rocco, M., Mirabelli-Badenier, M., Pedemonte, M., Panicucci, C., Trucks, H., Sander, T. *et al.* (2013) EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement. *J. Neurol.*, **260**, 1866–1870.
140. Le Duc, D., Horn, S., Jamra, R.A., Schaper, J., Wiczorek, D. and Redler, S. (2020) Novel EXOSC3 pathogenic variant results in a mild course of neurologic disease with cerebellum involvement. *Eur J Med Genet*, **63**, 103649.
141. Kasher, P.R., Namavar, Y., van Tijn, P., Fluiter, K., Sizarov, A., Kaimermans, M., Grierson, A.J., Zivkovic, D. and Baas, F. (2011) Impairment of the tRNA-splicing endonuclease subunit 54 (tsen54) gene causes neurological abnormalities and larval death in zebrafish models of pontocerebellar hypoplasia. *Hum. Mol. Genet.*, **20**, 1574–1584.
142. Namavar, Y., Chitayat, D., Barth, P.G., van Ruissen, F., de Wissel, M.B., Poll-The, B.T., Silver, R. and Baas, F. (2011) TSEN54 mutations cause pontocerebellar hypoplasia type 5. *Eur. J. Hum. Genet.*, **19**, 724–726.
143. Schaffer, A.E., Eggens, V.R., Caglayan, A.O., Reuter, M.S., Scott, E., Coufal, N.G., Silhavy, J.L., Xue, Y., Kayserili, H., Yasuno, K. *et al.* (2014) CLP1 founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. *Cell*, **157**, 651–663.
144. Lardelli, R.M., Schaffer, A.E., Eggens, V.R., Zaki, M.S., Grainger, S., Sathe, S., Van Nostrand, E.L., Schlachetki, Z., Rosti, B., Akizu, N. *et al.* (2017) Biallelic mutations in the 3' exonuclease TOE1 cause pontocerebellar hypoplasia and uncover a role in snRNA processing. *Nat. Genet.*, **49**, 457–464.
145. Li, Z., Schonberg, R., Guidugli, L., Johnson, A.K., Arnovitz, S., Yang, S., Scafidi, J., Summar, M.L., Vezina, G., Das, S. *et al.* (2015) A novel mutation in the promoter of RARS2 causes pontocerebellar hypoplasia in two siblings. *J. Hum. Genet.*, **60**, 363–369.
146. Giunta, M., Edvardson, S., Xu, Y., Schuelke, M., Gomez-Duran, A., Boczonadi, V., Elpeleg, O., Muller, J.S. and Horvath, R. (2016) Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy. *Hum. Mol. Genet.*, **25**, 2985–2996.
147. Gillespie, A., Gabunilas, J., Jen, J.C. and Chanfreau, G.F. (2017) Mutations of EXOSC3/Rrp40p associated with neurological diseases impact ribosomal RNA processing functions of the exosome in *S. cerevisiae*. *RNA*, **23**, 466–472.
148. Fasken, M.B., Losh, J.S., Leung, S.W., Brutus, S., Avin, B., Vaught, J.C., Potter-Birriell, J., Craig, T., Conn, G.L., Mills-Lujan, K. *et al.* (2017) Insight into the RNA Exosome Complex Through Modeling Pontocerebellar Hypoplasia Type 1b Disease Mutations in Yeast. *Genetics*, **205**, 221–237.
149. Ruggero, D. and Shimamura, A. (2014) Marrow failure: a window into ribosome biology. *Blood*, **124**, 2784–2792.
150. Kampen, K.R., Sulima, S.O., Vereecke, S. and De Keersmaecker, K. (2020) Hallmarks of ribosomopathies. *Nucleic Acids Res.*, **48**, 1013–1028.
151. Rocha, P.P., Micsinai, M., Kim, J.R., Hewitt, S.L., Souza, P.P., Trimarchi, T., Strino, F., Parisi, F., Kluger, Y. and Skok, J.A. (2012) Close proximity to Igh is a contributing factor to AID-mediated translocations. *Mol. Cell*, **47**, 873–885.
152. Weissbach, S., Langer, C., Puppe, B., Nedeva, T., Bach, E., Kull, M., Bargou, R., Einsele, H., Rosenwald, A., Knop, S. *et al.* (2015) The molecular spectrum and clinical impact of DIS3 mutations in multiple myeloma. *Br. J. Haematol.*, **169**, 57–70.
153. Walker, B.A., Wardell, C.P., Melchor, L., Hulkki, S., Potter, N.E., Johnson, D.C., Fenwick, K., Kozarewa, I., Gonzalez, D., Lord, C.J. *et al.* (2012) Intracloonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. *Blood*, **120**, 1077–1086.
154. Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J. *et al.* (2012) The origin and evolution of mutations in acute myeloid leukemia. *Cell*, **150**, 264–278.
155. Melchor, L., Brioli, A., Wardell, C.P., Murison, A., Potter, N.E., Kaiser, M.F., Fryer, R.A., Johnson, D.C., Begum, D.B., Hulkki Wilson, S. *et al.* (2014) Single-cell genetic analysis reveals the

- composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma. *Leukemia*, **28**, 1705–1715.
156. Walker, B.A., Boyle, E.M., Wardell, C.P., Murison, A., Begum, D.B., Dahir, N.M., Proszek, P.Z., Johnson, D.C., Kaiser, M.F., Melchor, L. *et al.* (2015) Mutational spectrum, copy number changes, and outcome: results of a sequencing study of patients with newly diagnosed myeloma. *J. Clin. Oncol.*, **33**, 3911–3920.
  157. Boyle, E.M., Ashby, C., Tytarenko, R.G., Deshpande, S., Wang, H., Wang, Y., Rosenthal, A., Sawyer, J., Tian, E., Flynt, E. *et al.* (2020) BRAF and DIS3 mutations associate with adverse outcome in a long-term follow-up of patients with multiple myeloma. *Clin. Cancer Res.*, **26**, 2422–2432.
  158. Lionetti, M., Barbieri, M., Todoerti, K., Agnelli, L., Fabris, S., Tonon, G., Segalla, S., Cifola, I., Pinal, E., Tassone, P. *et al.* (2015) A compendium of DIS3 mutations and associated transcriptional signatures in plasma cell dyscrasias. *Oncotarget*, **6**, 26129–26141.
  159. Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D. *et al.* (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*, **481**, 506–510.
  160. Chapman, M.A., Lawrence, M.S., Keats, J.J., Cibulskis, K., Sougnez, C., Schinzel, A.C., Harview, C.L., Brunet, J.P., Ahmann, G.J., Adli, M. *et al.* (2011) Initial genome sequencing and analysis of multiple myeloma. *Nature*, **471**, 467–472.
  161. Tomecki, R., Drazkowska, K., Kucinski, I., Stodus, K., Szczesny, R.J., Gruchota, J., Owczarek, E.P., Kalisiak, K. and Dziembowski, A. (2014) Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target. *Nucleic. Acids. Res.*, **42**, 1270–1290.
  162. Lohr, J.G., Stojanov, P., Carter, S.L., Cruz-Gordillo, P., Lawrence, M.S., Auclair, D., Sougnez, C., Knoechel, B., Gould, J., Saksena, G. *et al.* (2014) Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell*, **25**, 91–101.
  163. de Groen, F.L., Krijgsman, O., Tijssen, M., Vriend, L.E., Ylstra, B., Hooijberg, E., Meijer, G.A., Steenbergen, R.D. and Carvalho, B. (2014) Gene-dosage dependent overexpression at the 13q amplicon identifies DIS3 as candidate oncogene in colorectal cancer progression. *Genes Chromosomes Cancer*, **53**, 339–348.
  164. Pertesi, M., Vallee, M., Wei, X., Revuelta, M.V., Galia, P., Demangel, D., Oliver, J., Foll, M., Chen, S., Perriat, E. *et al.* (2019) Exome sequencing identifies germline variants in DIS3 in familial multiple myeloma. *Leukemia*, **33**, 2324–2330.
  165. Consortium, E.P., Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shores, N., Adrian, J., Kawli, T., Davis, C.A., Dobin, A. *et al.* (2020) Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, **583**, 699–710.
  166. Consortium, E.P., Snyder, M.P., Gingeras, T.R., Moore, J.E., Weng, Z., Gerstein, M.B., Ren, B., Hardison, R.C., Stamatoyannopoulos, J.A., Graveley, B.R. *et al.* (2020) Perspectives on ENCODE. *Nature*, **583**, 693–698.