PORTLAND PRESS

## Review Article

# The role of TREX in gene expression and disease

Catherine G. Heath, Nicolas Viphakone and Stuart A. Wilson

Department of Molecular Biology and Biotechnology, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, U.K.

Correspondence: Stuart A. Wilson (stuart.wilson@sheffield.ac.uk)

OPEN ACCESS

TRanscription and EXport (TREX) is a conserved multisubunit complex essential for embryogenesis, organogenesis and cellular differentiation throughout life. By linking transcription, mRNA processing and export together, it exerts a physiologically vital role in the gene expression pathway. In addition, this complex prevents DNA damage and regulates the cell cycle by ensuring optimal gene expression. As the extent of TREX activity in viral infections, amyotrophic lateral sclerosis and cancer emerges, the need for a greater understanding of TREX function becomes evident. A complete elucidation of the composition, function and interactions of the complex will provide the framework for understanding the molecular basis for a variety of diseases. This review details the known composition of TREX, how it is regulated and its cellular functions with an emphasis on mammalian systems.

## Introduction

Protein-coding genes transcribed by RNA polymerase II (Pol II) give rise to RNA molecules termed pre-mRNAs. During and/or after their synthesis, these precursors undergo a series of three main processing steps. Their 5′ end receives an $m^7G$ cap structure, shortly after the initiation of transcription, which protects the nascent pre-mRNA from degradation and plays a role in further stages of mRNA maturation [1]. The introns are excised during splicing which generally occurs co-transcriptionally, although approximately 20% of splicing events take place following transcription [2]. The 3′-end of the pre-mRNA is processed by endonucleolytic cleavage and polyadenylation, leaving the transcript with a polyadenosine (poly(A)) tail [3]. Following these processing events, the mRNA is exported from the nucleus through the nuclear pore complex to the cytoplasm where it can be translated into proteins (Figure 1).

From its site of synthesis until it reaches the nuclear pore, the future mRNA associates with numerous proteins to form a messenger ribonucleoprotein (mRNP). Throughout its nuclear odyssey, the mRNP composition is modified by the various remodeling events that accompany each maturation stage (Figure 1). The completion of each processing step results not only in a modified RNA molecule but also in the association of a specific set of proteins with the RNA. These two hallmarks of a mature mRNA are monitored by quality control mechanisms that trigger the degradation of defective mRNAs either in the nucleus by the nuclear exosome [4] or in the cytoplasm via the NMD pathway [5]. The expression of a functional protein thus relies on a properly synthesized and processed mRNA transcript packaged with the appropriate group of proteins.

A key player in mRNP biogenesis and maturation is the TRanscription and EXport (TREX) complex. TREX is conserved across a wide range of organisms including *Saccharomyces cerevisiae*, *Drosophila*, *Arabidopsis*, *Xenopus* and humans [6–10], indicating its key physiological importance. Components of TREX are functionally linked to each mRNA processing step, and this is reflected by their physical association with protein complexes such as the 5′ cap-binding complex (CBC), the exon junction complex (EJC) and 3′-end processing factors [11–13] (Figure 1). Therefore, TREX probably acts as an interface for these various processes to help maintain high fidelity of the gene expression process. In this review, focused primarily on mammalian TREX, we present the known composition with an emphasis on its dynamic organization uncovered over the last few years. We also summarize the current knowledge regarding how it functions in gene expression. Finally, we will present new fields in which TREX seems to have an essential role.
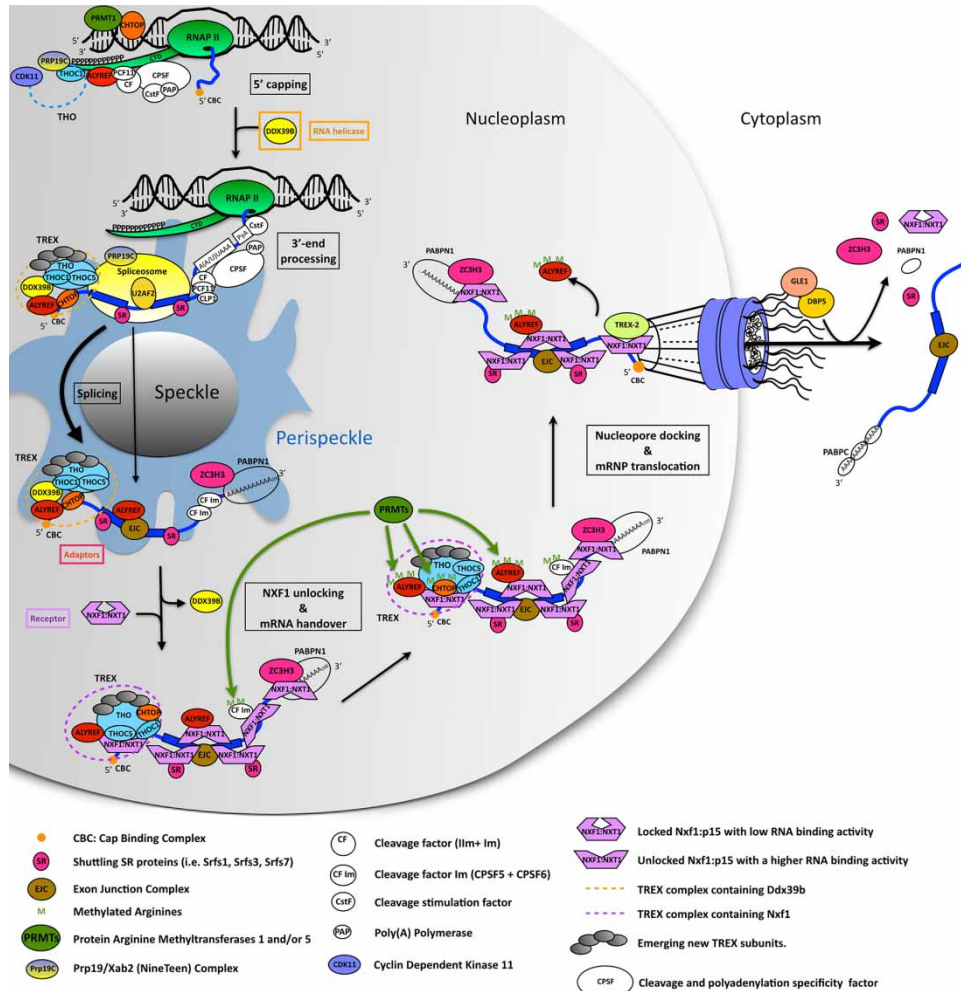
**Figure 1. Overview of NXF1- and TREX-dependent mRNP formation and nucleocytoplasmic transport.**

During gene expression, the TREX complex is recruited co-transcriptionally by a direct interaction with the phosphorylated C-terminal domain (CTD) of the large subunit of RNA polymerase II and via the PRP19:U2AF2 complex. Each processing step undergone by the pre-mRNA (5′-capping, co- and posttranscriptional splicing, 3′-end processing) acts as a trigger for TREX assembly and deposition along the transcript. Once bound to the mRNA, the combined action of adaptor (red and pink) and co-adaptor proteins (e.g. THOC5, CHTOP and CPSF6) recruit the NXF1:NXT1 heterodimer and cause a conformation change in NXF1, allowing exposure of its RBD for interaction with mRNA. In turn, the adaptor protein hands the mRNA over to NXF1. This process is promoted by protein arginine methyltransferase 1 (PRMT1) through arginine methylation of both the co-adaptor CHTOP, facilitating its interaction with NXF1, and the adaptor ALYREF, which reduces its RNA-binding activity enabling RNA handover to NXF1. Additional adaptor proteins, such as SR proteins and ZC3H3, can also recruit NXF1 to mRNA. Subsequently, direct interactions of NXF1 with the TREX-2 complex and nucleoporins allow the mRNP to dock at the nuclear pore, thus promoting export of the mRNP. At the nuclear pore export, adaptors are released from the mRNP, this is predicted to revert NXF1 to a low-affinity RNA-binding state that primes it for release from the mRNP. DBP5 and GLE1 trigger the release of NXF1 from the mRNP on the cytoplasmic side of the nuclear pore for recycling back to the nucleus.

# The composition of TREX

TREX consists of a stoichiometric salt-resistant hexameric core called THO, consisting of THOC1, 2, 5–7 and Tex1 [10] together with a group of additional proteins. These can come together in different combinations to make alternative forms of TREX [14–19] (Figure 2 and Table 1). The THO subunits were originally linked to mRNP biogenesis and export through their mutant phenotypes in *S. cerevisiae* [31,47,48]. Subsequently, their existence as a stable THO complex was discovered by their ability to suppress the transcriptional defects of the Hpr1 mutant by overexpression [34]. The genetic and physical interaction of THO with the DEAD-box RNA
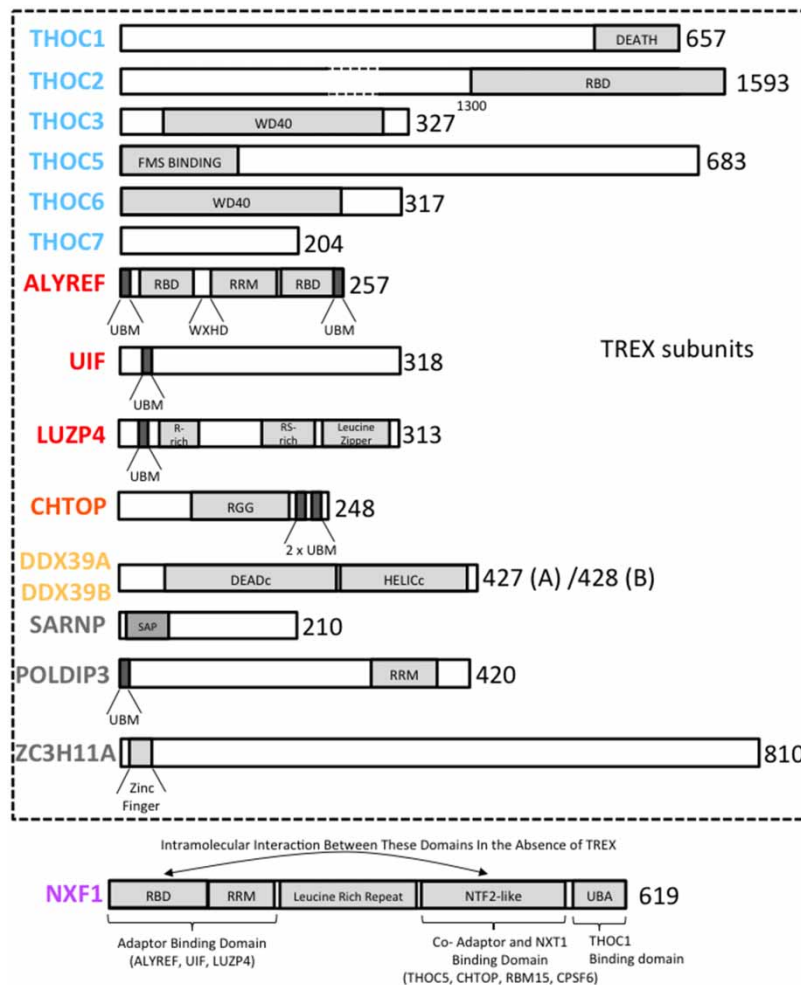
**Figure 2. Schematics of the domain structures for mRNA export factors.**

The major characterized domains within TREX subunits and NXF1 are shown. Each protein and domain is drawn approximately to scale and the size of each protein in amino acids is shown on the right hand side. RBD, RNA-binding domain; R-rich, arginine-rich; RS-rich, arginine/serine-rich; RRM, RNA recognition motif (not necessarily involved in RNA binding). UBM, UAP56 (DDX39B)-binding motif; WxHD corresponds to the recently identified motif which binds EIF4AIII [12]. The other domains are all well characterized. Note that THOC2 is not drawn to scale due to its size.

helicase Sub2p and the RNA export adaptor Yra1p and equivalent biochemical interactions among the human orthologs led to the discovery of the conserved TREX complex [6]. Additional mammalian TREX subunits have been identified through molecular associations and functional characterization through depletion and overexpression effects [10,14,17,24]. TREX subunits have discrete functions with diverse cellular roles, coming together to form this influential complex in the cell and the *curriculum vitae* for each subunit is presented below.

## THO

Human THOC1 is the ortholog of the yeast protein Hpr1 [6,15] and was originally discovered by its association with the tumor suppressor retinoblastoma protein (pRB) [20]. THOC1 is involved in cell cycle regulation and inducing p53-independent apoptosis through its death domain (Figure 2) [49–52]. THOC2 is the human ortholog of the yeast Tho2 protein and, as the largest subunit of TREX, is proposed to act as a scaffold for the formation of the complex [53]. THOC2 is an essential component for maintaining TREX function in humans [15,42], and mutations in the *THOC2* gene are associated with intellectual disabilities [22,54]. The conserved THOC3 (hTEX1) protein has been recently established as a component of the THO subcomplex in TREX and contains WD40 repeat motifs that allow multiple protein interactions (Figure 2) [10,15]. THOC5, 6 and 7 were

**Table 1 Composition of the TREX complex**

| | Human name | Alternative names | Cellular function | Yeast orthologs | kDa | Functional features | Key references | | | Link to BioGRID |
|---|---|---|---|---|---|---|---|---|---|---|
| THOsubunits | THOC1 | hHpr1, p84, p84N5, N5 | Apoptosis regulator | Hpr1p | 76 | Death domain, Rb-interacting factor | [20] | [21] | [196] | BIOGRID ENTRY |
| | THOC2 | hRlr1 | THO scaffold protein | Tho2p, Rlr1p | 183 | Coiled-coil domains, lysine-rich region | [31] | [15] | [22] | BIOGRID ENTRY |
| | THOC3 | Tex1, hTrex45 | | Tex1p | 39 | WD40 domains | [15] | [10] | | BIOGRID ENTRY |
| | THOC5 | fSAP79, Fmip | Export co-adaptor | | 79 | Leucine zipper, Fms-binding domain | [23] | [10] | [24] | BIOGRID ENTRY |
| | THOC6 | fSAP35, Wdr58 | | | 38 | WD40 domain | [10] | [15] | | BIOGRID ENTRY |
| | THOC7 | fSAP24, Nif3l1bp1 | | | 24 | Coiled-coil domain, Leucine zipper | [10] | | | BIOGRID ENTRY |
| TREXsubunits | DDX39B | Uap56, Bat1, P47 | Splicing factor, EJC-associated protein, RBP loading factor | Sub2p | 49 | ATP-dependent DEAD/DEAH-box RNA helicase | [25] | [6] | [26] | BIOGRID ENTRY |
| | DDX39A | Ddx39, Urh49 | Ddx39b paralogue, Putative EJC-associated protein | | 49 | ATP-dependent DEAD/DEAH-box RNA helicase | [27] | [28] | | BIOGRID ENTRY |
| | ALYREF | Aly, Ref, Bef, Thoc4 | Export adaptor, EJC-associated protein | Yra1p, She11p | 27 | RRM, Nxf1-binding, Ddx39-binding, UBM sequence | [29] | [6] | [30] | BIOGRID ENTRY |
| | UIF | Fyttd1 | Export adaptor, EJC-associated protein | | 37 | Functional homology with Alyref | [17] | [186] | | BIOGRID ENTRY |
| | LUZP4 | Hom-Tes-85, CT-8 | Export adaptor, EJC-associated protein | | 36 | Functional homology with Alyref, RS-dipeptides, Leucine zipper | [33] | [19] | | BIOGRID ENTRY |
| | CHTOP | Srag, C1orf77, Fop | Export co-adaptor, EJC-associated protein | | 26 | PRMT1-interaction domain, RGG box | [35] | [36] | [16] | BIOGRID ENTRY |
| | SARNP | Cip29, Tho1, Hcc-1 | | Tho1p | 24 | SAP domain | [37] | [14] | | BIOGRID ENTRY |
| | POLDIP3 | Skar, KIAA1649, PDIP46EJC-associated Protein | | | 46 | UBM-like sequence, RRM | [38] | [18] | | BIOGRID ENTRY |
| | ZC3H11A | Zc11a, ZC3HDC11A | | | 89 | Zinc finger protein | [14] | [18] | | BIOGRID ENTRY |
| | ERH | dDroer | | | 12 | Generally uncharacterized | [197] | [14] | [39] | BIOGRID ENTRY |

PORTLAND PRESS

**Table 1**   Continued

| | Human name | Alternative names | Cellular function | Yeast orthologs | kDa | Functional features | Key references | | | Link to BioGRID |
|---|---|---|---|---|---|---|---|---|---|---|
| TREX-associated proteins | NXF1 | Tap | RNA export receptor | Mex67p | 70 | Arginine-rich RBD, pseudo-RRM, NTF2-like domain, UBA | [40] | [41] | [42] | BIOGRID ENTRY |
| | NXT1 | p15 | Required for Nxf1 stabilization and mRNA export | Mtr2p | 15 | Stabilizes and binds to NXF1's NTF2-like domain | [198] | [199] | [43] | BIOGRID ENTRY |
| | ZC3H18 | Nhn1 | NEXT complex component | | 106 | CCCH-containing, Zinc-finger protein | [200] | [44] | [45] | BIOGRID ENTRY |
| | SRRT | Ars2 | pri-miRNAs processing, CBC effector in RNA 3′ processing | | 100 | Cap-binding protein | [201] | [202] | [203] | BIOGRID ENTRY |
| | NCBP3 | C17orf85, Elg | Involved in RNA export upon viral infection | | 71 | Cap-binding protein, forming an alternative CBC with Ncbp1 | [204] | [14] | [46] | BIOGRID ENTRY |
| | NCBP1 | Cbp80 | Transcription elongation, RNA export, RNA stability | Cbc1p | 71 | Cap-binding protein, forming the CBC with Ncbp2 | [205] | [11] | [15] | BIOGRID ENTRY |

**PORTLAND** PRESS

originally identified as part of THO in human and *Drosophila* cells, but have no known yeast orthologs [7,10,24]. Mammalian THOC5 is a cytoplasmic substrate for the macrophage-colony stimulating factor receptor termed FMS [23] and is directly involved in the differentiation of macrophages [55], and a role in hematopoiesis and cancer has been established [56]. The THO complex is implicated in embryonic stem (ES) cell self-renewal, being responsible for the export of pluripotency transcripts in mouse ES cells [57]. THOC7 is devoid of a nuclear localization sequence but forms a tight complex with THOC5, allowing the heterodimer to be translocated to the nucleus [9]. THOC3 and THOC6 contain WD40 repeat motifs (Figure 2) commonly used as protein interaction domains, which may contribute to THO assembly.

## DDX39B

DEAD-box protein 39B (DDX39B), widely known as U2AF65-associated protein 56 (UAP56), is the metazoan ortholog of the yeast Sub2p protein. First discovered as an essential factor for pre-mRNA splicing through its association with U2AF2 [25], this RNA-stimulated ATPase and DEAD-box RNA helicase promotes spliceosome assembly [26,58]. DDX39B is present on pre-mRNAs during splicing [15,59] and is required for the subsequent recruitment of the TREX subunit ALYREF to both spliced and intronless mRNAs [59,60]. On binding ALYREF, the RNA bound to DDX39B is transferred to ALYREF [16]. This is a part of an orchestrated sequence of RNA handover events that occur in nuclear mRNA export, whereby the mRNA acts like a baton in a relay race. DDX39B accompanies the mRNP to the nuclear pore, but is eventually displaced before nucleocytoplasmic translocation [61]. The displacement is triggered by recruitment of the mRNA export receptor NXF1, which binds to ALYREF in a mutually exclusive manner with DDX39B (Figure 1) [30]. In mammalian cells, a paralog of DDX39B exists, DDX39A, which has some functional redundancy with DDX39B, because depletion of both helicases is required to efficiently block mRNA export [17,27]. However, different groups of mRNAs have their export affected following knockdown of DDX39A or B, indicating that the two helicases most likely assemble into subtly different TREX complexes acting preferentially on specific groups of mRNAs [28].

## ALYREF

ALYREF was originally described as a transcriptional coactivator of the T-cell receptor α gene by binding LEF-1 and acute myeloid leukemia (AML)-1 transcription factors to form an enhancer-stimulating complex [29]. It was subsequently characterized as a molecular chaperone capable of stimulating the dimerization of basic leucine-zipper transcription factors, thereby inducing their DNA-binding activity. These studies also showed that ALYREF has a propensity to form an oligomeric complex *in vitro* with a molecular weight of 460 kDa corresponding to 14 ALYREF monomers [62]. ALYREF contains N- and C-terminal transient helices, known as the UAP56-binding motif (UBM), which are necessary and sufficient for interaction with DDX39A/B [17] (Figure 2). It also has a central RNA recognition motif (RRM), which binds RNA weakly through loops 1 and 4 [63]. The RRM is flanked by two unstructured arginine-rich regions that are the principal RNA-binding sites. Such separate RNA-binding domains set within unstructured flexible peptides [63] would provide ALYREF with the potential to bridge interactions between separate parts of an mRNA, which is consistent with the RNA annealing activity observed for its yeast ortholog, Yra1 [64]. Such an activity suggests that ALYREF may play an important role in packaging of the mRNP, which adopts a compact crescent-shaped structure during translocation from the site of synthesis to the nuclear pore [65]. The arginine-rich regions are also the principal sites for interaction with the mRNA export receptor NXF1, and binding of NXF1 to ALYREF triggers transfer of RNA from ALYREF to NXF1 [30]. Arginines within the RNA-binding region of ALYREF are methylated. These modifications have been shown to reduce the ability of ALYREF to bind RNA which is important for the handover of RNA from ALYREF to NXF1 [66]. ALYREF binds phosphoinositides *in vitro* and mutations in the N- and C-terminal arginine-rich regions alter this binding. However, whether this association is physiologically important remains unclear since the mutations that alter phosphoinositide binding, leading to altered ALYREF localization *in vivo*, may also have altered its RNA-binding activity [67]. Like DDX39A/B, ALYREF is displaced from the mRNP before nucleocytoplasmic translocation [61,68]. ALYREF is involved in the export of both spliced and intronless mRNAs [69], and RNA interference (RNAi)-mediated knockdown of ALYREF leads to disruption of mRNA export for around 3900 mRNAs in 293T cells [70]. Functional redundancy between ALYREF and other mRNA export factors such as UIF (described below) may account for the relatively small number of mRNAs whose export is affected following ALYREF RNAi.

## UIF

UIF (also known as FYTTD1) was first identified using a BLAST search with the ALYREF UBM and, although it harbors this motif, it has relatively little additional homology with ALYREF and other mammalian proteins outside this peptide. UIF functions in a redundant manner with ALYREF and knockdown of both proteins leads to a strong mRNA export block in mammalian cells. Knockdown of ALYREF also leads to a large up-regulation of UIF levels in cells, which probably accounts for the modest block in mRNA export seen following ALYREF RNAi [17]. Although UIF associates with ALYREF, this interaction is RNA-dependent, whereas the association with other TREX subunits is not [17]. Thus, UIF may form an alternative TREX complex without ALYREF and the RNA-dependent association with ALYREF may arise because multiple TREX complexes are loaded onto an mRNA.

## LUZP4

LUZP4 was originally characterized as a cancer/testis antigen [33], which represents a group of poorly characterized proteins whose expression is normally restricted to testis but are frequently up-regulated in cancer cells. Subsequently, a BLAST search revealed that LUZP4 also carries a UBM (Figure 2) and acts as an mRNA export factor associating with TREX subunits and complementing ALYREF RNAi *in vivo*. While the expression of LUZP4 is normally restricted to testis, it is up-regulated in melanoma cancer cells where it is required for their survival [19].

## CHTOP

CHTOP was originally identified as an RNA-binding protein that regulates the cell cycle [71]. It was subsequently shown to be a binding partner for protein arginine methyltransferase 1. It is involved in the expression of the γ-globin gene and is recruited to the estrogen target gene, pS2, in breast cancer cells [35,72]. Its role in gene expression is further shown by its presence in the 'five friends of methylated CHTOP' desumoylation complex that is recruited to the Zbp-89 transcription factor during transcriptional activation [36]. The direct function of CHTOP in transcriptional regulation is akin to that of ALYREF. Therefore, CHTOP may also be key in linking the stages of transcription, pre-mRNA processing and export as a component of TREX. CHTOP contains two copies of the UBM found in ALYREF, UIF and LUZP4 (Figure 2) and is capable of stimulating the ATPase and helicase activities of DDX39A/B. Both ALYREF and CHTOP are dependent on DDX39A/B for efficient loading onto mRNA *in vivo* [16].

## SARNP

SARNP (also known as CIP29) was first identified as an EPO-stimulated cytokine-induced protein involved in cell cycle progression [37]. The small protein contains a SAP DNA-binding motif (Figure 2), suggesting a possible role in direct transcriptional regulation and it was subsequently identified as a TREX component [14]. Its yeast ortholog, Tho1, was identified alongside Tho2 as suppressing transcriptional defects of *hpr1* mutants [31], forming the initial connection to TREX [31,73]. SARNP forms an ATP-dependent trimeric complex with DDX39A/B and ALYREF [14]. Moreover, SARNP stimulates the ATPase and helicase activities of DDX39A/B [16,74]. The plant ortholog of SARNP is known as MOS11 where it has also been shown to facilitate mRNA export [75].

## POLDIP3

POLDIP3 was identified as an S6 kinase substrate with homology to ALYREF, which contributes to cell growth regulation [38]. It localizes to nuclear speckles, but also enhances translation [76], so is likely to play a role in many stages of mRNP biogenesis. It associates with DDX39A/B and TREX in an ATP-dependent manner and its over expression leads to retention of poly(A)$^+$ RNAs in nuclear speckles [18].

## ZC3H11A

ZC3H11A co-purifies with multiple other TREX subunits [14] and recent work indicates that, similar to POLDIP3, it associates with TREX and DDX39A/B in an ATP-dependent manner. Knockdown of ZC3H11A by RNAi leads to a strong nuclear accumulation of mRNA [18]. Together, these data suggest that ZC3H11A is a bone fide TREX subunit, required for efficient mRNA export.

## Other proteins associated with TREX

Using immunoprecipitation with antibodies to THOC2, SARNP and DDX39A/B, a group of core proteins that associate with all three TREX subunits were identified using mass spectrometry. These included CBP80, CBP20, SRRT and NCBP3 (ELG), which all associate with the 5′ cap on mRNA and ERH [14]. ERH has multiple roles in the cell including roles in splicing and the cell cycle (reviewed in ref. [39]) and is a binding partner for POLDIP3 [77] and this may account for its association with TREX.

## Differences in TREX between species

The two major systems which have been used to study TREX are yeast and mammalian. While TREX is clearly conserved between these organisms (Table 1), there are also some important differences. For example, there are no known yeast orthologs for CHTOP, POLDIP3 and ZC3H11A. There is also divergence among the THO subunits. THOC1, 2 and 3 have orthologs in the two systems, but THOC5, 6 and 7 are specific to higher eukaryotes [7], with the yeast THO complex having two alternative subunits, Mtf2 and Thp2 [6]. Humans have a single ALYREF gene, whereas mice and yeast have two orthologous genes. UIF is conserved in mammals, but is absent from yeast and other higher eukaryotes such as *Drosophila* and *Caenorhabditis elegans*. LUZP4 is conserved in vertebrates but absent in other metazoans. These important species differences, with an expansion of TREX subunits during evolution, probably reflect the diverse biological roles that TREX plays particularly in multicellular organisms.

## Recruitment of TREX to mRNA through transcription and RNA processing

### Transcription

As a dynamic complex, TREX components are recruited to the developing mRNP at various stages of biogenesis. Chromatin remodeling and subsequent transcription form the first phase of mRNA generation. UIF is recruited to the transcription complex and the nascent RNA by association with the SSRP1 subunit of FACT [17], a histone chaperone and transcription elongation factor that remodels the H2A:H2B histone dimer, so RNA Pol II can transcribe the gene [78]. ALYREF is recruited by another histone chaperone Spt6, which works alongside FACT to remodel H3:H4 associated nucleosomes during transcription [79]. Spt6 binds the carboxy-terminal domain (CTD) of the large subunit of RNA Pol II and recruits IWS1. IWS1 in turn recruits ALYREF to actively transcribing genes [80].

The yeast TREX complex associates directly with the Pol II CTD; however, the interaction is partially dependent on the presence of nascent RNA [81,82]. This mechanism may ensure that TREX only associates with transcribing Pol II. Ser2 phosphorylation of the Pol II CTD is required for TREX binding and its occupancy on genes mirrors Ser2 phosphorylation, increasing in a 5′ to 3′ direction. When the increased recruitment of TREX towards the 3′-end of a gene is impaired, this specifically hinders the expression of long genes [81]. As TREX assembles on nascent RNA, there is a possibility that the CTD could become depleted of TREX subunits during transcription and this would be particularly acute on long genes. Therefore, Ser2-dependent enhanced recruitment of TREX to the CTD may ensure that it does not become depleted of TREX subunits during transcription of long genes. Interestingly, Mex67, the yeast ortholog of NXF1, is also necessary for Hpr1 (THOC1) stabilization during transcription [83], demonstrating the feedback and coordination between transcription and export steps.

The Prp19 complex, which functions in splicing, transcription and transcription-coupled DNA damage repair [84], is also required for efficient recruitment of TREX to transcribed genes in yeast. A mutation in the Syf1 subunit of the Prp19 complex results in less Hpr1, Sub2 and Yra1 associated with genes, especially at the 3′-end, suggesting that Prp19 acts to recruit TREX co-transcriptionally [85]. The human PRP19 complex associates with U2AF2 and together they promote Pol II CTD-dependent splicing activation [86]. Moreover, the PRP19 complex associates with TREX in human cells [14]. U2AF2, which directly binds DDX39A/B, may provide the link between TREX and the PRP19 complex, facilitating co-transcriptional recruitment of TREX to transcribed genes. An increasingly recurrent theme in gene expression is the coupling of each stage and reciprocal relationships between each step. It appears that this is also the case for transcription and export because TREX is required for efficient transcription elongation, particularly of long genes [21,87]. ALYREF affects transcription of a subset of genes [70], suggesting a reciprocal coupling between transcription and export via ALYREF, as it has been observed for other steps in the gene expression pathway [88].

## 5′ Cap formation

The NCBP1 (CBP80) and NCBP2 (CBP20) heterodimer forms the CBC, which associates with the 5′ cap. TREX is recruited to the 5′ cap by direct NCBP1–ALYREF and NCBP1–THO interactions [11,15]. Interestingly, SRRT also co-purifies with TREX [14]. SRRT associates with the CBC, the nuclear exosome targeting complex (NEXT) and together with ZC3H18 forms the CBC–NEXT complex [45]. ZC3H18 has independently been shown to recruit TREX to the hepatitis B virus post-transcriptional regulatory element and promote export of intronless mRNAs [44]. NEXT consists of three subunits, hMTR4, RBM7 and ZCCHC8, and multiple TREX subunits (ALYREF, CHTOP, THOC1,2,3,5,7, POLDIP3, UIF and DDX39A) co-purify with hMTR4 [45]. This raises the possibility that TREX bound to MTR4 might regulate the assembly of CBC–NEXT and the recruitment of the exosome. In this way, the assembly of TREX on the 5′ cap may be a critical decision point regarding whether an mRNA is degraded or exported during nuclear quality control. NCBP3 has recently been characterized as an alternative binding partner for NCBP1 at the cap, which can replace NCBP2. NCBP3 knockdown leads to a clear nuclear accumulation of mRNA, establishing its role in mRNA export. It also appears to act preferentially during times of stress such as during viral infection [46]. The earlier observation that NCBP3 associates with TREX [14] may account for the role of NCBP3 in mRNA export, by providing the means to ensure TREX recruitment to the 5′ end of mRNAs and promoting efficient mRNA export in times of stress. The length of an RNA determines its export pathway and RNAs shorter than 300 bases generally do not utilize the TREX:NXF1 pathway. This choice is determined early during transcription when hnRNP C associates with the 5′-end of the nascent RNA if it is long enough. hnRNP C binding to the 5′-end of the nascent RNA prevents recruitment of alternative export factors such as PHAX. This directs the RNA to the mRNA export pathway, whereby TREX probably displaces hnRNP C from the cap proximal region [89].

## Splicing

It has long been known that the presence on an intron enhances gene expression from an otherwise intronless cDNA and that splicing enhances mRNA export [90]. Using *in vitro* splicing reactions, it was shown that recruitment of human TREX is splicing-dependent [10,59]. The splicing dependence may be linked to the fact that DDX39A/B is involved in spliceosome assembly and additionally binds U2AF2, which recognizes the polypyrimidine tract within introns, early during spliceosome assembly. U2AF2 also co-operates with DDX39A/B to regulate nuclear retention of pre-mRNA [91]. Therefore, rearrangements of the U2AF2:DDX39A/B axis during splicing might couple TREX assembly on spliced mRNA with release of pre-mRNA retention signals. Interestingly, intron status is significant in yeast, despite direct co-transcriptional recruitment, as Yra1 is recruited preferentially to spliced mRNA [92]

The ability of TREX to be loaded onto spliced mRNA *in vitro* occurs in the absence of transcription and 3′-end cleavage/polyadenylation [10]. Furthermore, approximately 80% of splicing occurs co-transcriptionally in human cells [2], and so a considerable amount of TREX may be loaded onto mRNAs prior to 3′-end processing. Co-transcriptional splicing is thought to occur largely in regions of decompacted chromatin at the periphery of, or within, nuclear speckles. Moreover, TREX subunits interact *in vivo* on the periphery of nuclear speckles [93], suggesting that this is a major site for co-transcriptional TREX assembly. In contrast, post-transcriptional splicing appears to be largely restricted to nuclear speckles [2] and post-splicing release of mRNAs from speckles is dependent on TREX subunits [94].

Following splicing, the EJC is loaded onto the mRNA at a canonical position 24 bases upstream of the splice junction [95], furthermore the EJC co-purifies with TREX subunits [96,97]. These data led to the suggestion that the EJC provides a binding platform for mRNA export factors including the TREX subunit ALYREF and the mRNA export receptor NXF1 [96]. More recently, a specific sequence motif has been identified within ALYREF (WxHD) (Figure 2) that facilitates its interaction with the EJC subunit EIF4AIII. Thus, the EJC and CBC provide a stable binding platform for TREX on the 5′-end of mRNAs *in vitro* [12]. The involvement of the EJC in TREX loading is consistent with the recent observation that the EJC assembles on mRNAs in regions surrounding the nuclear speckles, termed perispeckles [98] (Figure 1), where TREX subunits also interact [93]. Whether mammalian TREX is loaded onto internal mRNA sites around EJCs *in vivo* remains to be seen. In yeast, this appears to be the case, with TREX subunits bound along the body of mRNAs [92,99].

Despite the central importance of splicing for normal mRNA export, when splicing is inhibited by drugs such as spliceostatin, this leads to the leakage of pre-mRNA to the cytoplasm [100]. However, the overall

amount exported is very small and the vast majority of unspliced pre-mRNA is retained in the nucleus when observed by fluorescence *in situ* hybridization [101]. Nevertheless, since some pre-mRNA does leak to the cytoplasm, this may indicate that export factors can be loaded onto unspliced pre-mRNA *in vivo*. Thus, one of the key effects of splicing in mRNA export may be to ensure nuclear retention of pre-mRNA already loaded with export factors.

While intron removal is important for the export of spliced mRNAs, TREX subunits also associate with intronless mRNAs [69] and the association between CBC and TREX is important for their export [102]. Interestingly, U2AF2 also promotes intronless mRNA export in *Drosophila* [103], which may be connected with its ability to bind DDX39B. Specific internal RNA elements have been identified which promote intronless mRNA stability and export [104,105]. These elements bind U2AF2, PRP19 complex and TREX subunits and, therefore, may act as surrogates for splicing on intronless mRNAs. The recruitment of TREX to internal sites on mRNAs may be important to ensure the binding of multiple copies of the export receptor NXF1, ensuring smooth passage of the elongated mRNP through the nuclear pore. While TREX can function as a binding platform for NXF1, some intronless mRNAs appear to utilize additional strategies to ensure NXF1 recruitment. For example, histone mRNAs are bound by SR proteins that can in turn directly recruit NXF1 in a manner similar to ALYREF [106–108].

### 3′-end processing

A necessary pre-requisite for mRNA export from the nucleus to the cytoplasm is the release of mRNA from the DNA template, which is coupled to pre-mRNA 3′-end processing [3]. In yeast, two key proteins, Pcf11 and Clp1 come together to cause pre-mRNA cleavage. Pcf11 is recruited early during transcription through the Pol II CTD and also recruits Yra1 at the 3′-end of genes. Clp1 recruitment to Pcf11 triggers displacement of Yra1 from Pcf11 [13]. The ability of Yra1 to regulate accessibility of Clp1 to Pcf11 means that it can influence the site of cleavage/polyadenylation within a pre-mRNA, particularly at genes with divergent efficiency elements [109]. The displacement of Yra1 from Pcf11 by Clp1 is enhanced by the presence of Sub2, suggesting that the process is coordinated with loading Yra1 onto mRNA. Intriguingly, in both *tho* and *sub2* mutants, a stalled intermediate in mRNP biogenesis is created in which nuclear pore components and polyadenylation factors remain associated with chromatin in what is known as a 'heavy chromatin' fraction [110]. Together, these data indicate that TREX plays a key role in 3′-end processing and the concomitant release of the mRNP from chromatin.

The PCF11-ALYREF interaction is conserved in humans [13] and DDX39B has been found to be present in a highly purified 3′-end processing complex [111]. Thus, poly(A) site choice in mammalian cells is also likely to be governed by ALYREF, CLP1, DDX39A/B and PCF11, through similar mechanisms to those used in yeast. Additionally, THOC5 binds two 3′-end processing factors, CPSF100 [112] and CPSF6 [113]. Through these interactions, THOC5 can regulate poly(A) site choice for specific genes. In the case of CPSF6, THOC5 directs its early recruitment to transcribed genes and, following THOC5 or CPSF6 loss, proximal poly(A) sites are preferentially used [113]. A further twist in the 'tail' arises from recent studies on CDK11 that triggers Ser2 phosphorylation of the Pol II CTD [114]. Ser2 phosphorylation peaks at the 3′-end of genes and this stimulates the recruitment of 3′-end processing factors to the gene. TREX forms a stable complex with CDK11 and is required for its recruitment to the 3′-end of HIV 1 genes and subsequent CTD Ser2 phosphorylation. Therefore, TREX orchestrates 3′-end processing via CDK11 on HIV transcripts, but the extent to which TREX works with CDK11 on cellular mRNAs remains to be seen. In summary, TREX acts at many points in the evolution of the 3′-end of mRNA through interactions with multiple components of the transcription and 3′-end processing machinery. In turn, 3′-end processing, a necessary pre-requisite for mRNA export, no doubt allows the evolution of TREX to a state where it can subsequently trigger export of the fully processed mRNA.

## Integrating mRNA processing with TREX assembly

The early assembly of TREX subunits probably occurs on the RNA Pol II CTD through multiple mechanisms, including direct binding of subunits to the Ser2 phosphorylated CTD and indirectly via the PRP19:U2AF2 complex. Subsequently, TREX subunits transfer to the nascent mRNP. This process is driven in part by the RNA helicases DDX39A and B, which load export factors such as UIF, ALYREF, LUZP4 and CHTOP onto the mRNP, using the helicase ATPase cycle [14,16] (Figure 3). The assembly on the mRNP is driven by the major pre-mRNA processing events (capping, splicing and polyadenylation) that provide their own protein signatures

such as the CBC and EJC. These protein signatures probably serve to stabilize TREX at multiple positions along the mRNP (Figure 1). Interestingly, subunits of the THO complex are largely devoid of RNA-binding activity in yeast with the exception of Tho2 [53], and in mammals, THOC5 binds RNA weakly [42]. In contrast DDX39A/B, ALYREF and CHTOP bind RNA avidly [16] and these may well provide the main connection between TREX and the mRNA. Whether all ALYREF and CHTOP molecules associated with the mRNP are also bound by THO is not clear. Analysis of the relative amounts of mRNA export and RNA processing factors in mouse cells shows that DDX39B, SARNP, ALYREF and CHTOP are present in considerable excess over THO subunits and the RNA Pol II large subunit (Figure 4). For example, there is 90 times more DDX39B than THOC5 in the cell [115]. While ALYREF, CHTOP, SARNP and DDX39B are all present in excess of the cap-binding protein NCBP1, THO subunits are present in substoichiometric amounts, with the most abundant THO subunit, THOC7 being present at a ratio of approximately 1:2 with NCBP1. These variations in the levels of TREX subunits in the cell may reflect differential recruitment of specific TREX factors to subsets of mRNAs. In favor of this hypothesis, depletion of THOC5 alters the export of approximately 2.9% of mRNAs in mouse cells and depletion of multiple THO subunits in *Drosophila* only effects export of approximately 20% of the transcriptome [7]. However, the restricted sets of mRNAs affected could also be caused by functional redundancy between export factors such as that seen between ALYREF and UIF [17]. THO may alternatively play its major role in export by chaperoning the recruitment of ALYREF, DDX39B, SARNP and CHTOP onto mRNA through its association with RNA polymerase II. Consistent with this model, THO levels are much more comparable with RNA Pol II large subunit levels (Figure 4) and in yeast, chromatin immunoprecipitation experiments reveal that THO associates with the majority of transcribed genes [116]. The two possibilities are not mutually exclusive, but what is clear is that there are likely to be subtly different forms of TREX associating on different mRNAs and within a single mRNP, there are likely to be varying flavors of TREX present. The >3 fold excess of DDX39B, ALYREF and SARNP, which form a stable ATP-dependent trimer [14] over the nuclear cap-binding proteins (Figure 4), indicate that the cell has the opportunity to recruit multiple copies of this trimer to a single mRNP. Thus, these proteins may play an important role in mRNP packaging as well as export, consistent with the RNA annealing activity of Yra1 (ALYREF) [64]. Finally, it is striking that PCF11 is such a low abundance protein relative to THO or other mRNA export factors, with approximately 1000 times more ALYREF present in the cell. This vast excess of ALYREF probably helps ensure that PCF11 remains saturated with ALYREF until 3′-end processing, thus guarding against inappropriate poly(A) site choice. The substoichiometric amounts of PCF11 with respect to RNA polymerase II may indicate that it is used on selective transcripts in mammalian cells.

# Biological roles of TREX
## mRNA export
TREX plays a central role in mRNA export and this activity is governed by its ability to act as a binding platform for NXF1 [42]. Nuclear export factor 1 (NXF1, also called tip-associated protein, TAP) was originally identified in a yeast two-hybrid screen to identify cellular partners of tyrosine kinase-interacting protein (Tip), a viral protein from *Herpesvirus saimiri* [117]. A link between NXF1 and mRNA export was revealed by a synergistic combination of yeast genetic studies and biochemical experiments performed in metazoan systems. The yeast protein Mex67p was first shown to be essential for bulk poly(A)$^+$ RNA nuclear export, consistent with its nuclear localization, its ability to bind mRNAs *in vivo* and to contact the nuclear pore complex [118]. Concurrently, NXF1 was shown to promote nuclear export of retroviral transcripts by binding to RNA secondary structures called constitutive transport elements (CTEs) [119]. Subsequently, it was shown that Mex67p's function in mRNA export is conserved in metazoans through its ortholog NXF1 [40].

NXF1 is a multidomain protein (Figure 2), with each domain connected by flexible linkers. Extensive structural characterization has revealed an unstructured N-terminal region preceding the following folded domains: an RRM fold, a leucine-rich repeat, an NTF2-like (NTF2L) domain and finally a C-terminal UBA domain [43,120,121]. The NTF2L domain forms a tight complex with NXT1, which is essential for the stability and activity of NXF1 [43]. The NTF2L and UBA domains provide separate binding sites for FG repeat sequences found in the nucleoporins lining the nuclear pore. These FG repeat binding sites help direct NXF1 to the nuclear pore [43].

An additional conserved protein complex known as TREX-2, which consists of GANP, PCID2, DSS1 and ENY2 in mammals, is involved in the docking of NXF1 at the nuclear pore during mRNA export (Figure 1, [122,123]).
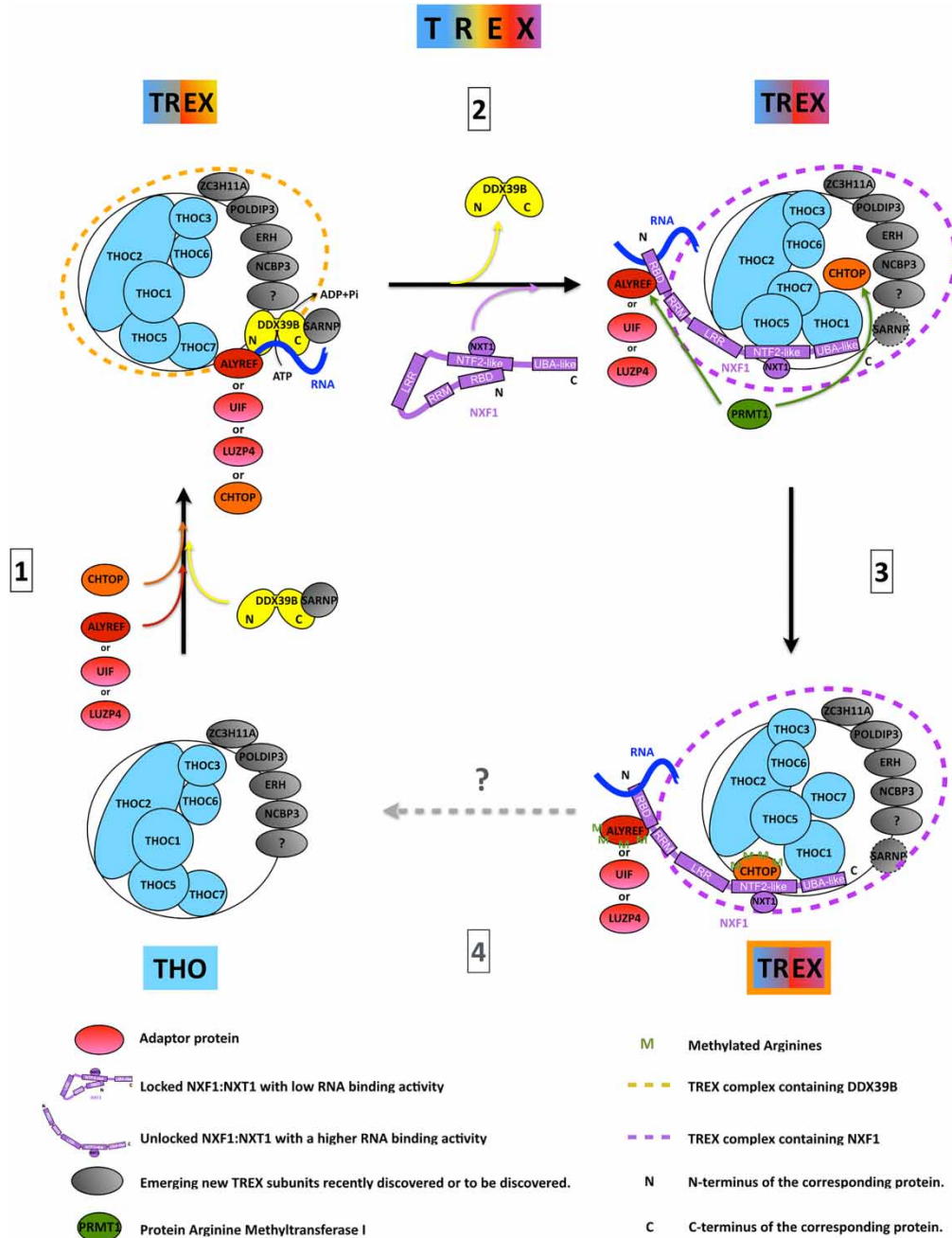
**Figure 3. The life cycle of the TREX complex.**

The THO complex constitutes a salt-resistant core of tightly associated proteins (light blue). During gene expression, it dynamically associates with a variety of proteins (red, orange, yellow, purple and gray) to form the TREX complex, whose composition evolves during mRNP formation. THO is able to recruit the adaptor proteins (red), the co-adaptor CHTOP and the RNA helicase DDX39B. The DDX39B RNA helicase is thought to use rounds of ATP hydrolysis to load adaptors and co-adaptors onto the RNA. In turn, they recruit the mRNA export receptor NXF1:p15, which displaces DDX39B. While the co-adaptor THOC5, as a part of THO, is likely to associate early with the NTF2L domain of NXF1, arginine methylation of the second major co-adaptor CHTOP is a prerequisite for it to bind that same domain. This suggests that rearrangements occur within TREX. It is currently unclear whether the additional subunits (gray) are all part of the same TREX complex or belong to variants of a remodeled TREX complex. It is also unknown whether the same THO complex is recycled for further rounds of mRNP assembly (putative step 4).
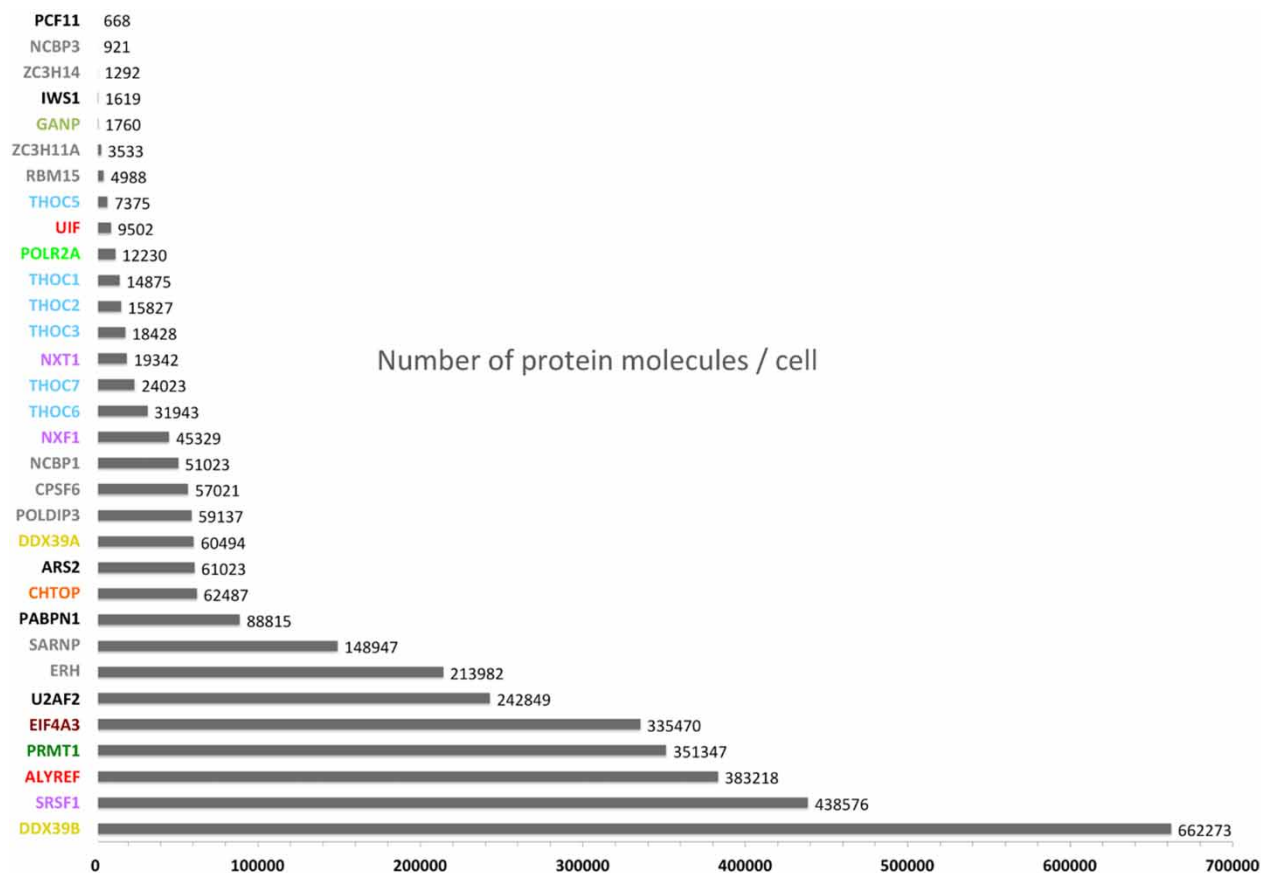
**Figure 4. The number of protein molecules per cell for mRNA export and processing factors.**
This graph shows the average number of protein molecules per cell for a mouse NIH3T3 fibroblast cell line. The data were extracted from ref. [115].

Certain subunits of TREX-2 are stably associated with the nuclear pore, but whether this association requires on-going transcription is currently unclear since conflicting data exist in the literature [122,124]. Work on the yeast complex has shown that it is involved in the translocation of genes to the nuclear pore and provides a direct connection between a gene's promoter, transcription and the nuclear pore through the mediator complex [125]. Furthermore, the ENY2 subunit of TREX-2 is also a component of the SAGA transcription activation complex [126], and thus TREX-2 may use multiple connections with the transcription machinery to direct a gene to the nuclear pore. Studies in the human system have shown that GANP is involved in the selective export of specific mRNAs [127,128], particularly those involved in gene expression steps, including RNA processing, splicing and ribosome biogenesis. GANP uses its FG repeat-like domain to interact with the FG repeat-binding regions of NXF1 and stimulates a local high concentration of NXF1 at the nuclear pore [122]. Therefore, TREX-2 may provide a fast track for the export of certain transcripts by positioning the gene close to the nuclear pore and promoting loading of NXF1 onto the mRNP and subsequent export. The use of TREX-2 by certain genes may be important for cells to respond rapidly to cellular stresses [128].

A systematic study of the RNA-binding activity of NXF1 domains using UV cross-linking revealed that the major RNA-binding activity resides within the N-terminal unstructured region amino acids 61–118, with weaker RNA-binding activity identified within amino acids 1–60 and the RRM (amino acids 119–198) [30]. These studies are corroborated by *in vivo* mRNP capture assays, which mapped the mRNP-binding site for NXF1 to amino acids 61–140 [129]. The presence of an RRM has caused some confusion in the literature over the years regarding the NXF1 mRNA-binding activity. In fact, the major RNA-binding domain (RBD) corresponds to amino acids 1–118, not the RRM (amino acids 118–198). Moreover, the LRR, NTF2L and UBA domains have no RNA-binding activity when assayed by UV cross-linking [30], though weak RNA-binding

activity has been detected for the NTF2L domain using electrophoretic mobility shift assays [130]. The NXF1 RBD is rich in arginines and mutation of 10 of these arginines prevents NXF1 mRNA binding both *in vitro* and *in vivo*. Moreover, a mutant form of NXF1, which carries the 10 arginines mutated to alanine, fails to function in mRNA export *in vivo* [30], establishing the RBD as the major determinant for mRNA binding. The binding of NXF1 to retroviral CTE elements appears to be fundamentally different from binding to mRNA and involves extensive interactions between the RRM, LRR and NTF2L domains of NXF1 [41,121,130]. Interestingly, the *NXF1* gene harbors a CTE-like structure within intron 10 and NXF1 binding to this sequence promotes the export of a retained intron transcript that encodes a short form of NXF1, though the function of this short form of NXF1 is unclear [131].

ALYREF binds to the N-terminal region of NXF1, including the RBD and RRM domains [132] (Figure 2). As ALYREF binds mRNA avidly, it was given the name mRNA export adaptor, because it was thought to bridge the interaction between RNA and NXF1. However, subsequent studies showed that when NXF1 binds to an ALYREF:RNA complex, the RNA is handed over to NXF1 [30] and ALYREF no longer binds RNA. This is consistent with biochemical studies, showing that the NXF1-binding site and RNA-binding sites on ALYREF overlap [63]. Additional mRNA export adaptors have been identified including various SR proteins, which, like ALYREF, use an arginine-rich peptide to bind the RBD + RRM domains of NXF1 [106–108]. A common feature of the RNA handover process from an export adaptor to NXF1 is that once the adaptor protein is bound to NXF1 it enhances NXF1 RNA binding activity [30]. The reasons for this were unclear until it was realized that NXF1 forms an intramolecular interaction between the RBD and the NTF2L domains which suppresses RNA binding (Figure 3) [42]. Thus, the NTF2L domain of NXF1 is autoinhibitory for mRNA binding driven by the RBD. ALYREF is involved in disrupting the NXF1 intramolecular interaction, which exposes the RBD for interaction with RNA. A second site for interaction between TREX and NXF1 involves the NXF1 UBA domain that binds THOC1 [42]. The third site is the NTF2L domain, which THOC5 binds on the opposite side to that bound by NXT1. The binding of ALYREF to NXF1 stimulates the binding of THOC5 to the complex; on the basis of this activity, THOC5 was named a co-adaptor for mRNA export [24]. The NXF1 RBD binds to the NTF2L domain at a site overlapping that bound by THOC5 and it is the combined action of ALYREF and THOC5 binding to NXF1 that exposes its RBD, allowing optimal RNA-binding activity [42]. Since NXF1 requires both an adaptor and a co-adaptor for optimal RNA binding, this raises the question of whether SR proteins function independently to recruit NXF1 to mRNAs, given they would only provide the adaptor function. Evidence from yeast suggests that SR proteins may function in collaboration with TREX during mRNA export [133–135], and this may be to ensure that a co-adaptor is available at such sites for optimal NXF1 RNA binding.

THOC5 is not the only co-adaptor protein that binds the NTF2L domain of NXF1; CHTOP, RBM15 and CPSF6 also bind this domain [16,136,137]. Similar to THOC5, CHTOP binding to NXF1 is stimulated by the presence of ALYREF [16]. CHTOP and THOC5 are found in the same TREX complex bound to NXF1; therefore, there is potentially sequential binding of these factors to NXF1 during mRNP maturation (Figure 3). The methylation of arginines within CHTOP is essential for its interaction with NXF1 [16] and methylation of arginines within ALYREF is required for RNA handover to NXF1 [66]. Therefore, arginine methylation plays a vital role in the activity of TREX during mRNA export.

An interesting feature of two of the known NXF1 co-adaptors, THOC5 and CPSF6, is their involvement in 3′-end processing and polyadenylation. Furthermore, NXF1 loss in cells leads to hyperadenylation of mRNA and loss of Mex67 in yeast leads to retention of the 3′-end processing factors Rna14 and Rna15 on the mRNA [138]. A further factor involved in the recruitment of NXF1 to the mRNP is ZC3H3 and loss of this factor leads to mRNA hyperadenylation, similar to that seen following NXF1 depletion [139]. These observations suggest that an adaptor, such as ALYREF, delivered via PCF11 and a co-adaptor, such as THOC5, delivered through interactions with 3′-end processing factors provide a key signal at the 3′-end of mRNAs. This signal is interpreted by NXF1 that subsequently binds to TREX, allowing correct evolution of the 3′-end processing complex to a state suitable for mRNA export.

Once NXF1 has been delivered to the mRNP by TREX, the mRNP is escorted to the nuclear pore [128]. Studies on *Chironomus tentans* have shown that ALYREF and DDX39B dissociate from the mRNP on the nuclear side or during translocation through the nuclear pore [61]. This is predicted to revert NXF1 to a low-affinity RNA-binding state and this may be an important prerequisite for its subsequent removal from the mRNP. Dbp5, alongside its ATPase activator Gle1 and inositol hexakisphosphate phosphate 6, work together with NUP159 to remodel the mRNP on the cytoplasmic side of the nuclear pore, triggering removal of NXF1 [140–142]. NXF1 is then reimported into the nucleus for subsequent rounds of mRNA export.

PORTLAND
PRESS

## piRNA biogenesis

While the majority of work on TREX has thus far focused on its role in mRNA export, recent studies indicate that it has additional roles in piRNA biogenesis. piRNAs are short RNA molecules produced in germ line cells with a major function in suppressing transposons during development. The first indication that TREX may be involved in piRNA biogenesis came from the observation that DDX39A/B co-localized at piRNA clusters with Rhino, which is required for transposon silencing [143]. Moreover, DDX39A/B mutations disrupted a cytoplasmic structure known as the nuage, which is rich in components of the piRNA processing machinery. Most recently, it has been shown that additional TREX subunits are involved in piRNA transcription and biogenesis, with TREX loading onto piRNAs being driven by a chromatin-associated protein, Cutoff [144]. However, it is not clear yet whether TREX plays a role in the transport of piRNA precursors from the nucleus to the cytoplasm.

## Genome stability, R-loops and cancer

During transcription, the nascent RNA can hybridize with the single-stranded template DNA present within the transcription bubble. This structure is known as an R-loop and it leads to the opposing DNA strand being left single-stranded, making it susceptible to cleavage which can eventually lead to various forms of DNA damage in cells [145,146]. TREX is intimately associated with both the RNA Pol II CTD and nascent RNA, leaving it poised to play a major role in sequestering the nascent RNA and preventing it from hybridizing to the template DNA strand. A large increase in R-loops is found in TREX-depleted yeast and human cells leading to far more DNA damage, transcription-associated recombination and DNA replication obstacles [87,116,147,148]. R-loops also form part of a natural process of genomic recombination to generate variation, such as for antibody class-switching in B-cells. THO depletion in murine cells results in an increased rate of class-switching as more R-loops are formed and consequently, a greater rate of recombination can occur [87]. An RNA/DNA helicase, SETX (senataxin; Sen1 in yeast), is required to resolve R-loop formation. Interestingly in yeast, Sen1 co-purifies with Yra1 (ALYREF) [149], raising the possibility that as well as packaging RNA to help prevent R-loop formation, ALYREF might also be involved in the regulation of SETX activity. A further role for TREX in maintaining genome stability arises from studies on IPMK, which is involved in inositol phosphate production in human cells. Loss of this kinase prevents the export of mRNAs encoding proteins involved in the DNA damage response. ALYREF is implicated in this process and thus, TREX maintains appropriate levels of proteins required to maintain genome stability by ensuring efficient export of their mRNAs [150]. Since TREX can associate with various different adaptor proteins such as UIF and LUZP4 (Figure 3), this raises the possibility that alternative forms of TREX promote the export of specific classes of mRNA [128].

Given the importance of TREX in maintaining genome stability, it is unsurprising that TREX has been implicated in many forms of cancer. CHTOP is recruited to the pS2 promoter, inducing the expression of estrogen target genes in breast cancer cells [35,36] and thereby maintaining growth of the tumor. CHTOP is also required for the expression of key genes required for the maintenance of tumor cells in glioblastoma [151]. THOC1 associates with the tumor suppressor pRB, resulting in the prevention of cell cycle arrest and THOC1-induced apoptosis [20,49,50]. This links THOC1 with tumor survival as mRNP biogenesis cannot occur efficiently enough to maintain rapid cell division in its absence [152,153]. The up-regulation of THOC1 and THO complex expression, readily seen in a range of cancer types, may correlate with tumor size and have a greater effect in hormone-dysregulated cancers [152–155]. Phosphorylation of THOC5 on tyrosine 225 promotes its incorporation into mRNPs and THOC5 phosphorylation by leukemogenic protein tyrosine kinases is increased in patients with chronic myeloid leukemia [156]. ALYREF also shows an altered expression pattern in various cancerous tissues, but is more frequently up-regulated [154,157]. Depletion of ALYREF results in the reduced metastatic capacity of human oral squamous cell carcinoma cell lines, demonstrating the increased TREX dependence in cancer cells. Interestingly, ALYREF depletion in this cell line results in increased levels of 'metastasis modulating molecules' RRP1B and CD82 that suppress the metastatic capacity of tumor cells [158–160]. ALYREF directly interacts with RRP1B, so the increased expression of ALYREF in cancer cells may titrate away RRP1B allowing metastasis to develop uninterrupted [157]. DDX39B depletion affects the expression levels of the tumor suppressor BRCA1 protein, which is important for the DNA damage response [28]. SARNP is up-regulated in cancer tissues [37,161], but a further link to AML has been made attributable to a translocation event creating an SARNP–mixed lineage leukemia (MLL) fusion protein [162]. As MLL does not have any intrinsic DNA-binding capability,

the SARNP–MLL fusion may confer this action through SARNP's SAP DNA-binding motif (Figure 2), potentially resulting in altered gene expression, a hallmark of cancer development [163].

TREX seems to be directly involved in regulating the cell cycle, consistent with a role in proliferation and differentiation, and this activity probably correlates with the aberrant expression of TREX subunits in cancer cells. During mitosis, the cell moves through characteristic stages and TREX may function at the intermediate checkpoints. THOC1 and THOC2 influence the rate of proliferation, with THOC1 playing a more specific role in the G2-M checkpoint [50,54]. ALYREF is significant in the S and G2 phases where its expression seems co-ordinated with that of Cyclin A [67]. ALYREF also represses the activity of the E2F2 transcription factor when bound to the CHK1 promoter, preventing the expression of the CHK1 checkpoint kinase that induces cell cycle arrest and checkpoint activation [164]. Furthermore, DDX39B and THOC2 may play a role in chromosome alignment [165] and spindle assembly [166], though whether these are direct effects or indirect through failure to export crucial mRNAs remains unclear.

The numerous associations of TREX in cancer development could lead to new specific therapeutics. The accumulating evidence that cancerous tissues are more dependent on TREX to aid their survival and progression suggests that drugs to deplete TREX levels and/or activity could be a potential target for cancer therapy. The marked reduction in metastatic capacity and cancer cell proliferation following ALYREF [157], THOC1 [21,153] and LUZP4 [19] depletion further suggests that there may be therapeutic benefit in targeting TREX.

## Cellular differentiation

In addition to elevated TREX dependence in cancerous tissue, ES cells show a greater susceptibility to TREX loss. ES cells are pluripotent and derived from the inner cell mass (ICM) of the pre-implantation embryo at the blastocyst stage. Depletion of THO components at this stage results in embryonic lethality as the blastocyst cannot be effectively formed due to increased apoptosis of the ICM [56,57,167]. However, depletion of these components in differentiated adult cells does not result in immediate cell death [168] and THOC1 expression is decreased in nonproliferating cells [52], suggesting a greater dependency on TREX in undifferentiated cells early on in embryogenesis. Furthermore, the balance between differentiation and self-renewal within stem cells may be regulated by TREX through its function in mRNA export. Reduced levels of THOC2 and THOC5 during differentiation result in a decreased expression of self-renewal factors NANOG and SOX2, but an increased expression of differentiation markers GATA3 and BRACHURY [57,169]. THOC2 and THOC5 associate with differentiation marker mRNAs more readily than self-renewal mRNAs and the nuclear accumulation of NANOG and SOX2 mRNA suggests that this regulation occurs through mRNA export [57].

The differentiation and development of specific tissues show a greater demand for TREX. The levels of THOC5 and SARNP are greater in the human fetal liver, where hematopoietic cells differentiate [23], and SARNP is up-regulated in the adult bone marrow [37]. THOC5 is known to regulate differentiation of hematopoietic cells to various lineages [55] and may do so by regulating the expression of specific transcription factors involved in monocyte development [170]. The regulation of immediate early genes involved in macrophage differentiation is dependent on THOC5, perhaps linked to 3′-end processing of the transcripts [112]. Bone marrow and spleen cells are specifically affected following the loss of THOC5, resulting in death after a few days in mice [56,171], consistent with the loss of differentiating cells when THOC5 is depleted [172]. Conversely, THOC5 is dispensable in the terminally differentiated liver [172]. The small intestine is remarkably sensitive to the loss of TREX, presumably due to the presence of stem-cell niches in the intestinal crypt. THOC1 expression is higher in intestinal crypts harboring stem-cell niches [168], and THOC1 depletion results in degeneration of the small-intestinal epithelium and increased apoptosis of cells in the niche [168,172]. Interestingly, the stem-cell niches in the large intestine are not as affected by THOC1 depletion [168], demonstrating the tissue-specific effects of TREX loss as well as the increased requirement for the complex in differentiating tissues. Male gametogenesis requires rapid proliferation and stem-cell differentiation in the testes to produce sperm. Concurrent with increased TREX dependency in differentiating cells, fertility in male mice was greatly reduced when two hypomorphic THOC1 alleles were present [173]. Female mice also had reduced fertility when homozygous for the hypomorphic allele, most probably due to defects in the ICM during embryogenesis as mentioned above.

From studies in cancer, embryogenesis and tissue differentiation, the dependence on TREX for cellular survival seems to be increased in more rapidly proliferating and differentiating cells. Proliferating cells will have a higher rate of replication, transcription and therefore processing, export and translation to maintain cell

division and induce differentiation. This increased burden on the gene expression machinery probably dictates the dependency on TREX that is central to this pathway.

## Neurodevelopment and neurodegenerative diseases

The role of TREX in neuronal disorders can be divided into two groups: direct mutations in TREX subunits, which trigger neuronal disease, and mutations in other genes, which lead to a potential dependence on TREX during the disease state. In the first class, missense mutations in the *THOC2* gene cause syndromic intellectual disability [22]. Additionally, a *de novo* translocation event in the vicinity of *THOC2* on the X-chromosome has been described in a child with cerebellar hypoplasia, ataxia and retardation [54]. The translocation of the PTK2 kinase gene generates a fusion transcript with THOC2 and leads to a decrease in expression of both encoded proteins. As inactivation of PTK2 alone does not cause this phenotype, the effects on *THOC2* are likely to be hugely influential in causing this disease, corroborated by *C. elegans* knockout studies showing impaired movement [54]. *THOC6* mutations have also been associated with intellectual disability, consistent with its high expression levels in the developing zebrafish brain [174]. The THOC6 mutation results in a mislocalization of the protein to the cytoplasm, suggesting that its nuclear role as a part of TREX is essential for proper neural and organ development. Furthermore, a homozygous mutation in a conserved region of *THOC6* is the likely cause of a disease with clinical features such as intellectual disability, brain malformation and renal and heart defects [175].

A second class of neuronal disease that indirectly depends on TREX is amyotrophic lateral sclerosis (ALS), associated with death of motor neurons. The most common familial form of ALS involves a GGGGCC repeat expansion within an intron of the *C9ORF72* gene. Following transcription, the repeat expanded pre-mRNA is exported to the cytoplasm where it forms RNA foci. Thus, the transcript overrides the normal nuclear retention mechanisms that keep pre-mRNA in the nucleus. The repeat expansion can also be translated in multiple reading frames in an ATG start codon-independent manner, leading to the production of toxic dipeptides (reviewed in ref. [176,177]). A survey of proteins that bind the repeat expansion identified multiple proteins including ALYREF, leading to the suggestion that excess ALYREF recruitment to the repeat expanded pre-mRNA, may stimulate its export [178]. A *Drosophila* model for neurodegeneration triggered by the C9ORF72 was recently used to carry out a genetic screen for suppressors and enhancers of neurodegeneration [179]. This screen identified a mutation disrupting ALYREF activity as a potent suppressor of neurodegeneration in this system and intriguingly identified CHTOP and NXF1 as enhancers. Furthermore, it was shown that the toxic peptides that are produced by the repeat expansion block mRNA export in human cells [179]. Therefore, in C9ORF72 disease, one unexpected reason for cell death may be the inhibition of expression of multiple proteins through a general mRNA export block. This export block is triggered by the inappropriate export of the pre-mRNA containing the repeat and this may involve the excessive recruitment of ALYREF to the repeat expansion.

## Viral replication

Viruses replicate by utilizing host cell machineries, which are essential for them to infect more cells and organisms. One such machinery is the NXF1–NXT1 mRNA export pathway that is employed to export viral mRNA, under the guise of a host mRNA. This 'hijacking' is conserved among viral species and subfamilies as it is used in Influenza [32,180,181], various Herpes strains and the Hepatitis B virus. The mRNA export pathway is typically hijacked by the expression of viral-specific export adaptors that serve to recruit TREX and NXF1 to viral mRNA. The Kaposi's Sarcoma-associated Herpes Virus and Herpes virus saimiri ORF57 are the most well studied of these adaptors, recruiting ALYREF to the viral mRNA to which ORF57 is bound [182–185]. The rest of TREX is also recruited, so the established RNA handover to NXF1–NXT1 can occur [184,186]. The Herpes simplex virus ICP27 [185,187], Epstein-Barr virus EB2 [188] and Human cytomegalovirus UL69 [189] proteins are functionally homologous to ORF57 as they similarly recruit ALYREF, the rest of TREX and NXF1 for viral RNA export. In some cases, DDX39B seems to have more of an influence on viral RNA export than ALYREF [189,190], supporting the role of the whole TREX complex in viral mRNA export and efficient viral replication. In addition, ALYREF may act to stabilize viral mRNAs at the 3′-end [191]. As well as viral export adaptors, elements in viral RNA act to recruit the host cell mRNA export machinery. The posttranscriptional regulatory element is necessary for efficient expression of Hepatitis B proteins and triggers recruitment of TREX [44]. Constitutive transport elements have been identified in a number of simple retroviruses, including Mason Pfizer monkey and Simian retrovirus, which directly bind NXF1 with high affinity, promoting export of viral

transcripts [119]. Therefore, viruses have evolved a number of different strategies to hijack the cellular mRNA export machinery to ensure efficient expression of their own proteins.

## Plant development

Unsurprisingly, TREX seems to play a role in other RNA pathways. The existence and general function of TREX seems conserved in plants [8,75,192,193], and it also appears to play a role in the transport of small-interfering RNA precursors for long-distance gene-silencing effects [8,194]. siRNA interaction with AGO may require TREX to localize the siRNA, or to target the AGO protein necessary for gene silencing [8,194]. THO mutants disrupt these pathways differentially depending on the plant's developmental stage, consistent with human systems showing a greater dependence on TREX during embryogenesis and in differentiating tissues. TREX is implicated in other plant processes such as disease resistance and ethylene signaling, which is essential for development. Ethylene-induced senescence is increased in Hpr1 mutants, due to defective transcription of ethylene signaling and ribosomal protein genes [193,195]. The involvement of TREX in pathways other than mRNA export suggests that it may be a more general RNA trafficking module in the cell beyond plants, capable of influencing the localization of other RNA classes yet to be discovered.

# Discussion

The conserved TREX complex is highly integrated into cellular function and homeostasis (Figure 5), functioning throughout mRNA biogenesis and in many other cellular mechanisms. Its presence is necessary for effective transcription, processing and export as well as preventing DNA damage and maintaining differentiation during embryogenesis and in specific adult tissues. The dynamic nature of the proteins that associate with TREX is likely to provide the flexibility for the complex to carry out multiple functions and link mRNA generating and processing steps together. The less well-characterized members of TREX such as ERH, ZC3H11A and POLDIP3 are likely to reveal further functions of TREX and refine the functions already known. A key and important step in understanding mammalian TREX function in the future will be to elucidate its RNA substrates, transcriptome-wide, as has recently been done for yeast TREX [92,99]. The additional functions of TREX in the cell cycle, proliferating cells and the prevention of DNA damage must be investigated further, especially as understanding TREX could facilitate the development of novel drugs to treat cancer, ALS and viral infections.
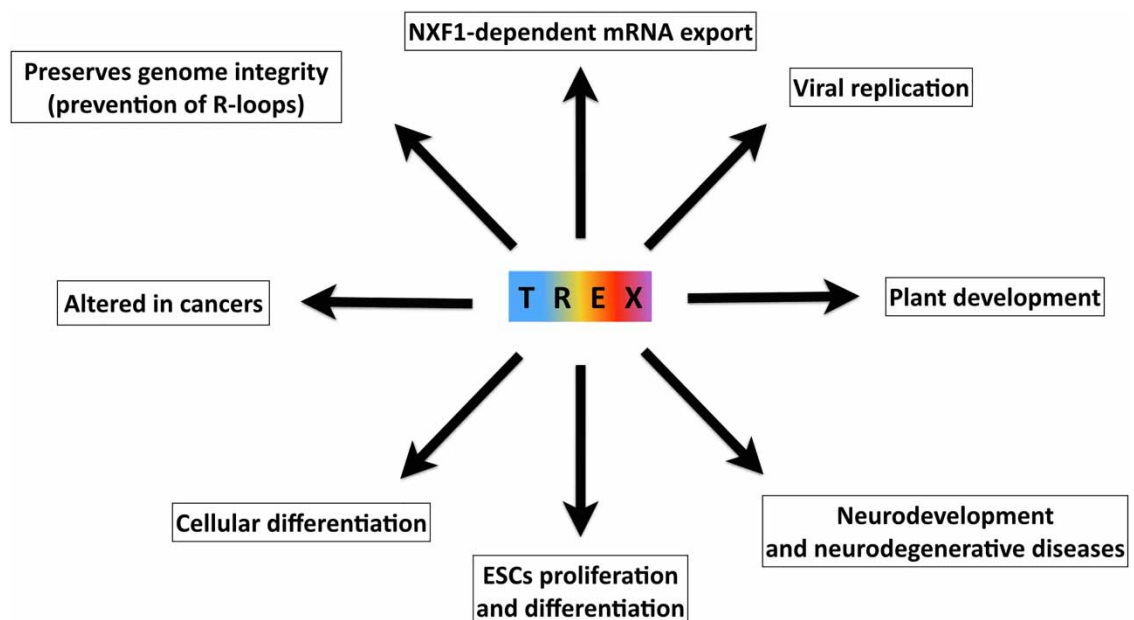


**Figure 5. An overview of the biological processes involving TREX.**
ESC: embryonic stem cell.

PORTLAND
PRESS

## Abbreviations

ALS, amyotrophic lateral sclerosis; AML, acute myeloid leukemia; CBC, cap-binding complex; CTD, carboxy terminal domain; CTEs, constitutive transport elements; DDX39B, DEAD-box protein 39B; EJC, exon junction complex; ES, embryonic stem; EPO, erythropoietin; ICM, inner cell mass; MLL, mixed lineage leukemia; mRNP, messenger ribonucleoprotein; NEXT, nuclear exosome targeting complex; NMD, nonsense mediated decay; NTF2L, NTF2-like; NXF1, nuclear export factor 1; Pol II, polymerase II; poly(A), polyadenosine A; pRB, retinoblastoma protein; PRMT1, protein arginine methyltransferase 1; RBD, RNA-binding domain; RNAi, RNA interference; RRM, RNA recognition motif; TAP, tip-associated protein; TAR, transcription-associated recombination; Tip, tyrosine kinase interacting protein; TREX, TRanscription and EXport; UAP56, U2AF65-associated protein 56; UBM, UAP56-binding motif.

## Acknowledgements

## Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

## References

1　Topisirovic, I., Svitkin, Y.V., Sonenberg, N. and Shatkin, A.J. (2011) Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev.* **2**, 277–298 doi:10.1002/wrna.52

2　Girard, C., Will, C.L., Peng, J., Makarov, E.M., Kastner, B., Lemm, I. et al. (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat. Commun.* **3**, 994 doi:10.1038/ncomms1998

3　Proudfoot, N.J. (2011) Ending the message: poly(A) signals then and now. *Genes Dev.* **25**, 1770–1782 doi:10.1101/gad.17268411

4　Kilchert, C., Wittmann, S. and Vasiljeva, L. (2016) The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.* doi:10. 1038/nrm.2015.15

5　Trcek, T., Sato, H., Singer, R.H. and Maquat, L.E. (2013) Temporal and spatial characterization of nonsense-mediated mRNA decay. *Genes Dev.* **27**, 541–551 doi:10.1101/gad.209635.112

6　Strässer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S. et al. (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**, 304–308 doi:10.1038/nature746

7　Rehwinkel, J., Herold, A., Gari, K., Köcher, T., Rode, M., Ciccarelli, F.L. et al. (2004) Genome-wide analysis of mRNAs regulated by the THO complex in *Drosophila melanogaster. Nat. Struct. Mol. Biol.* **11**, 558–566 doi:10.1038/nsmb759

8　Yelina, N.E., Smith, L.M., Jones, A.M., Patel, K., Kelly, K.A. and Baulcombe, D.C. (2010) Putative Arabidopsis THO/TREX mRNA export complex is involved in transgene and endogenous siRNA biosynthesis. *Proc. Natl Acad. Sci. USA* **107**, 13948–13953 doi:10.1073/pnas.0911341107

9　El Bounkari, O., Guria, A., Klebba-Faerber, S., Claußen, M., Pieler, T., Griffiths, J.R. et al. (2009) Nuclear localization of the pre-mRNA associating protein THOC7 depends upon its direct interaction with Fms tyrosine kinase interacting protein (FMIP). *FEBS Lett.* **583**, 13–18 doi:10.1016/j.febslet. 2008.11.024

10　Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N. and Reed, R. (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* **19**, 1512–1517 doi:10.1101/gad.1302205

11　Cheng, H., Dufu, K., Lee, C.S., Hsu, J.L., Dias, A. and Reed, R. (2006) Human mRNA export machinery recruited to the 5′ end of mRNA. *Cell* **127**, 1389–1400 doi:10.1016/j.cell.2006.10.044

12　Gromadzka, A.M., Steckelberg, A.L., Singh, K.K., Hofmann, K. and Gehring, N.H. (2016) A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs. *Nucleic Acids Res.* doi:10.1093/nar/gkw009

13　Johnson, S.A., Cubberley, G. and Bentley, D.L. (2009) Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3′ end processing factor Pcf11. *Mol. Cell* **33**, 215–226 doi:10.1016/j.molcel.2008.12.007

14　Dufu, K., Livingstone, M.J., Seebacher, J., Gygi, S.P., Wilson, S.A. and Reed, R. (2010) ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev.* **24**, 2043–2053 doi:10.1101/gad.1898610

15　Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M. et al. (2013) Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res.* **41**, 1294–1306 doi:10.1093/nar/gks1188

16　Chang, C.-T., Hautbergue, G.M., Walsh, M.J., Viphakone, N., van Dijk, T.B., Philipsen, S. et al. (2013) Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J.* **32**, 473–486 doi:10.1038/emboj.2012.342

17　Hautbergue, G.M., Hung, M.L., Walsh, M.J., Snijders, A.P.L., Chang, C-T., Jones, R. et al. (2009) UIF, a new mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr. Biol.* **19**, 1918–1924 doi:10.1016/j.cub.2009.09.041

18　Folco, E.G., Lee, C.S., Dufu, K., Yamazaki, T. and Reed, R. (2012) The proteins PDIP3 and ZC11A associate with the human TREX complex in an ATP-dependent manner and function in mRNA export. *PLoS ONE* **7**, e43804 doi:10.1371/journal.pone.0043804

19　Viphakone, N., Cumberbatch, M.G., Livingstone, M.J., Heath, P.R., Dickman, M.J., Catto, J.W. et al. (2015) Luzp4 defines a new mRNA export pathway in cancer cells. *Nucleic Acids Res.* **43**, 2353–2366 doi:10.1093/nar/gkv070

20　Durfee, T., Mancini, M.A., Jones, D., Elledge, S.J. and Lee, W.H. (1994) The amino-terminal region of the retinoblastoma gene product binds a novel nuclear matrix protein that co-localizes to centers for RNA processing. *J. Cell Biol.* **127**, 609–622 doi:10.1083/jcb.127.3.609

21  Li, Y., Wang, X., Zhang, X. and Goodrich, D.W. (2005) Human hHpr1/p84/Thoc1 regulates transcriptional elongation and physically links RNA polymerase II and RNA processing factors. *Mol. Cell. Biol.* **25**, 4023–4033 doi:10.1128/MCB.25.10.4023-4033.2005

22  Kumar, R., Corbett, M.A., van Bon, B.W.M., Woenig, J.A., Weir, L., Douglas, E. et al. (2015) THOC2 mutations implicate mRNA-export pathway in X-Linked intellectual disability. *Am. J. Hum. Genet.* **97**, 302–310 doi:10.1016/j.ajhg.2015.05.021

23  Tamura, T., Mancini, A., Joos, H., Koch, A., Hakim, C., Dumanski, J. et al. (1999) FMIP, a novel Fms-interacting protein, affects granulocyte/macrophage differentiation. *Oncogene* **18**, 6488–6495 doi:10.1038/sj.onc.1203062

24  Katahira, J., Inoue, H., Hurt, E. and Yoneda, Y. (2009) Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *EMBO J.* **28**, 556–567 doi:10.1038/emboj.2009.5

25  Fleckner, J., Zhang, M., Valcarcel, J. and Green, M.R. (1997) U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev.* **11**, 1864–1872 doi:10.1101/gad.11.14.1864

26  Shen, J., Zhang, L. and Zhao, R. (2007) Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. *J. Biol. Chem.* **282**, 22544–22550 doi:10.1074/jbc.M702304200

27  Pryor, A., Tung, L., Yang, Z., Kapadia, F., Chang, T.H. and Johnson, L.F. (2004) Growth-regulated expression and G0-specific turnover of the mRNA that encodes URH49, a mammalian DExH/D box protein that is highly related to the mRNA export protein UAP56. *Nucleic Acids Res.* **32**, 1857–1865 doi:10.1093/nar/gkh347

28  Yamazaki, T., Fujiwara, N., Yukinaga, H., Ebisuya, M., Shiki, T., Kurihara, T. et al. (2010) The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Mol. Biol. Cell.* **21**, 2953–2965 doi:10.1091/mbc.E09-10-0913

29  Bruhn, L., Munnerlyn, A. and Grosschedl, R. (1997) ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRalpha enhancer function. *Genes Dev.* **11**, 640–653 doi:10.1101/gad.11.5.640

30  Hautbergue, G.M., Hung, M.-L., Golovanov, A.P., Lian, L.-Y. and Wilson, S.A. (2008) Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc. Natl Acad. Sci. USA* **105**, 5154–5159 doi:10.1073/pnas.0709167105

31  Piruat, J.I. and Aguilera, A. (1998) A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *EMBO J.* **17**, 4859–4872 doi:10.1093/emboj/17.16.4859

32  Hao, L., Sakurai, A., Watanabe, T., Sorensen, E., Nidom, C.A., Newton, M.A. et al. (2008) Drosophila RNAi screen identifies host genes important for influenza virus replication. *Nature* **454**, 890–893 doi:10.1038/nature07151

33  Türeci, O., Sahin, U., Koslowski, M., Buss, B., Bell, C., Ballweber, P. et al. (2002) A novel tumour associated leucine zipper protein targeting to sites of gene transcription and splicing. *Oncogene* **21**, 3879–3888 doi:10.1038/sj.onc.1205481

34  Chavez, S., Beilharz, T., Rondon, A.G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J.Q. et al. (2000) A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae. EMBO J.* **19**, 5824–5834 doi:10.1093/emboj/19.21.5824

35  van Dijk, T.B., Gillemans, N., Stein, C., Fanis, P., Demmers, J., van de Corput, M. et al. (2010) Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol. Cell Biol.* **30**, 260–272 doi:10.1128/MCB.00645-09

36  Fanis, P., Gillemans, N., Aghajanirefah, A., Pourfarzad, F., Demmers, J., Esteghamat, F. et al. (2012) Five friends of methylated chromatin target of protein-arginine-methyltransferase[prmt]-1 (chtop), a complex linking arginine methylation to desumoylation. *Mol. Cell Proteomics* doi:10.1074/mcp.M112.017194

37  Fukuda, S., Wu, D.W., Stark, K. and Pelus, L.M. (2002) Cloning and characterization of a proliferation-associated cytokine-inducible protein, CIP29. *Biochem. Biophys. Res. Commun.* **292**, 593–600 doi:10.1006/bbrc.2002.6680

38  Richardson, C.J., Bröenstrup, M., Fingar, D.C., Jülich, K., Ballif, B.A., Gygi, S. et al. (2004) SKAR is a specific target of S6 kinase 1 in cell growth control. *Curr. Biol.* **14**, 1540–1549 doi:10.1016/j.cub.2004.08.061

39  Weng, M.T. and Luo, J. (2013) The enigmatic ERH protein: its role in cell cycle, RNA splicing and cancer. *Protein Cell* **4**, 807–812 doi:10.1007/s13238-013-3056-3

40  Katahira, J., Sträßer, K., Podtelejnikov, A., Mann, M., Jung, J.U. and Hurt, E. (1999) The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J.* **18**, 2593–2609 doi:10.1093/emboj/18.9.2593

41  Teplova, M., Wohlbold, L., Khin, N.W., Izaurralde, E. and Patel, D.J. (2011) Structure-function studies of nucleocytoplasmic transport of retroviral genomic RNA by mRNA export factor TAP. *Nat. Struct. Mol. Biol.* **18**, 990–998 doi:10.1038/nsmb.2094

42  Viphakone, N., Hautbergue, G.M., Walsh, M., Chang, C.-T., Holland, A., Folco, E.G. et al. (2012) TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat. Commun.* **3**, 1006 doi:10.1038/ncomms2005

43  Fribourg, S., Braun, I.C., Izaurralde, E. and Conti, E. (2001) Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. *Mol. Cell.* **8**, 645–656 doi:10.1016/S1097-2765(01)00348-3

44  Chi, B., Wang, K., Du, Y., Gui, B., Chang, X., Wang, L. et al. (2014) A sub-element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. *Nucleic Acids Res.* **42**, 7305–7318 doi:10.1093/nar/gku350

45  Andersen, P.R., Domanski, M., Kristiansen, M.S., Storvall, H., Ntini, E., Verheggen, C. et al. (2013) The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat. Struct. Mol. Biol.* **20**, 1367–1376 doi:10.1038/nsmb.2703

46  Gebhardt, A., Habjan, M., Benda, C., Meiler, A., Haas, D.A., Hein, M.Y. et al. (2015) mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nat. Commun.* **6**, 8192 doi:10.1038/ncomms9192

47  Chavez, S. and Aguilera, A. (1997) The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes Dev.* **11**, 3459–3470 doi:10.1101/gad.11.24.3459

48  Schneiter, R., Guerra, C.E., Lampl, M., Gogg, G., Kohlwein, S.D. and Klein, H.L. (1999) The *Saccharomyces cerevisiae* hyperrecombination mutant hpr1 Δ is synthetically lethal with two conditional alleles of the acetyl coenzyme A carboxylase gene and causes a defect in nuclear export of polyadenylated RNA. *Mol. Cell Biol.* **19**, 3415–3422 doi:10.1128/MCB.19.5.3415

49  Doostzadeh-Cizeron, J., Evans, R., Yin, S. and Goodrich, D.W. (1999) Apoptosis induced by the nuclear death domain protein p84N5 is inhibited by association with Rb protein. *Mol. Biol. Cell.* **10**, 3251–3261 doi:10.1091/mbc.10.10.3251

50  Doostzadeh-Cizeron, J., Terry, N.H. and Goodrich, D.W. (2001) The nuclear death domain protein p84N5 activates a G2/M cell cycle checkpoint prior to the onset of apoptosis. *J. Biol. Chem.* **276**, 1127–1132 doi:10.1074/jbc.M006944200

51  Evans, R.L., Poe, B.S. and Goodrich, D.W. (2002) Nuclear localization is required for induction of apoptotic cell death by the Rb-associated p84N5 death domain protein. *Oncogene* **21**, 4691–4695 doi:10.1038/sj.onc.1205583

52  Gasparri, F., Sola, F., Locatelli, G. and Muzio, M. (2004) The death domain protein p84N5, but not the short isoform p84N5s, is cell cycle-regulated and shuttles between the nucleus and the cytoplasm. *FEBS Lett.* **574**, 13–19 doi:10.1016/j.febslet.2004.07.074

53  Peña, A., Gewartowski, K., Mroczek, S., Cuéllar, J., Szykowska, A., Prokop, A. et al. (2012) Architecture and nucleic acids recognition mechanism of the THO complex, an mRNP assembly factor. *EMBO J.* **31**, 1605–1616 doi:10.1038/emboj.2012.10

54  Di Gregorio, E., Bianchi, F.T., Schiavi, A., Chiotto, A.M.A, Rolando, M., Verdun di Cantogno, L. et al. (2013) A de novo X;8 translocation creates a PTK2-THOC2 gene fusion with THOC2 expression knockdown in a patient with psychomotor retardation and congenital cerebellar hypoplasia. *J. Med. Genet.* **50**, 543–551 doi:10.1136/jmedgenet-2013-101542

55  Mancini, A., Koch, A., Whetton, A.D. and Tamura, T. (2004) The M-CSF receptor substrate and interacting protein FMIP is governed in its subcellular localization by protein kinase C-mediated phosphorylation, and thereby potentiates M-CSF-mediated differentiation. *Oncogene* **23**, 6581–6589 doi:10.1038/sj.onc.1207841

56  Mancini, A., Niemann-Seyde, S.C., Pankow, R., El Bounkari, O., Klebba-Färber, S., Koch, A. et al. (2010) THOC5/FMIP, an mRNA export TREX complex protein, is essential for hematopoietic primitive cell survival in vivo. *BMC Biol.* **8**, 1 doi:10.1186/1741-7007-8-1

57  Wang, L., Miao, Y.-L., Zheng, X., Lackford, B., Zhou, B., Han, L. et al. (2013) The THO complex regulates pluripotency gene mRNA export and controls embryonic stem cell self-renewal and somatic cell reprogramming. *Cell Stem Cell* **13**, 676–690 doi:10.1016/j.stem.2013.10.008

58  Shen, H., Zheng, X., Shen, J., Zhang, L., Zhao, R. and Green, M.R. (2008) Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. *Genes Dev.* **22**, 1796–1803 doi:10.1101/gad.1657308

59  Luo, M.-J., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M. et al. (2001) Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413**, 644–647 doi:10.1038/35098106

60  Taniguchi, I. and Ohno, M. (2008) ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Mol. Cell Biol.* **28**, 601–608 doi:10.1128/MCB.01341-07

61  Kiesler, E., Miralles, F. and Visa, N. (2002) HEL/UAP56 binds cotranscriptionally to the Balbiani ring pre-mRNA in an intron-independent manner and accompanies the BR mRNP to the nuclear pore. *Curr. Biol.* **12**, 859–862 doi:10.1016/S0960-9822(02)00840-0

62  Virbasius, C.M., Wagner, S. and Green, M.R. (1999) A human nuclear-localized chaperone that regulates dimerization, DNA binding, and transcriptional activity of bZIP proteins. *Mol. Cell.* **4**, 219–228 doi:10.1016/S1097-2765(00)80369-X

63  Golovanov, A.P., Hautbergue, G.M., Tintaru, A.M., Lian, L.-Y. and Wilson, S.A. (2006) The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. *RNA* **12**, 1933–1948 doi:10.1261/rna.212106

64  Portman, D.S., O'Connor, J.P. and Dreyfuss, G. (1997) YRA1, an essential *Saccharomyces cerevisiae* gene, encodes a novel nuclear protein with RNA annealing activity. *RNA* **3**, 527–537 PMID: 9149233

65  Björk, P. and Wieslander, L. (2015) The Balbiani Ring Story: synthesis, assembly, processing, and transport of specific messenger RNA-protein complexes. *Annu. Rev. Biochem.* **84**, 65–92 doi:10.1146/annurev-biochem-060614-034150

66  Hung, M.L., Hautbergue, G.M., Snijders, A.P., Dickman, M.J. and Wilson, S.A. (2010) Arginine methylation of REF/ALY promotes efficient handover of mRNA to TAP/NXF1. *Nucleic Acids Res.* **38**, 3351–3361 doi:10.1093/nar/gkq033

67  Okada, M., Jang, S.-W. and Ye, K. (2008) Akt phosphorylation and nuclear phosphoinositide association mediate mRNA export and cell proliferation activities by ALY. *Proc. Natl Acad. Sci. USA* **105**, 8649–8654 doi:10.1073/pnas.0802533105

68  Kim, V.N., Yong, J., Kataoka, N., Abel, L., Diem, M.D. and Dreyfuss, G. (2001) The Y14 protein communicates to the cytoplasm the position of exon-exon junctions. *EMBO J.* **20**, 2062–2068 doi:10.1093/emboj/20.8.2062

69  Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M. and Izaurralde, E. (2001) REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. *Proc. Natl Acad. Sci. USA* **98**, 1030–1035 doi:10.1073/pnas.98.3.1030

70  Stubbs, S.H. and Conrad, N.K. (2015) Depletion of REF/Aly alters gene expression and reduces RNA polymerase II occupancy. *Nucleic Acids Res.* **43**, 504–519 doi:10.1093/nar/gku1278

71  Zullo, A.J., Michaud, M., Zhang, W. and Grusby, M.J. (2009) Identification of the small protein rich in arginine and glycine (SRAG): a newly identified nucleolar protein that can regulate cell proliferation. *J. Biol. Chem.* **284**, 12504–12511 doi:10.1074/jbc.M809436200

72  van Dijk, T.B., Gillemans, N., Pourfarzad, F., van Lom, K., von Lindern, M., Grosveld, F. et al. (2010) Fetal globin expression is regulated by friend of Prmt1. *Blood* **116**, 4349–4352 doi:10.1182/blood-2010-03-274399

73  Jimeno, S., Luna, R., Garcia-Rubio, M. and Aguilera, A. (2006) Tho1, a novel hnRNP, and Sub2 provide alternative pathways for mRNP biogenesis in yeast THO mutants. *Mol. Cell Biol.* **26**, 4387–4398 doi:10.1128/MCB.00234-06

74  Sugiura, T., Sakurai, K. and Nagano, Y. (2007) Intracellular characterization of DDX39, a novel growth-associated RNA helicase. *Exp. Cell Res.* **313**, 782–790 doi:10.1016/j.yexcr.2006.11.014

75  Germain, H., Qu, N., Cheng, Y.T., Lee, E., Huang, Y., Dong, O.X. et al. (2010) MOS11: a new component in the mRNA export pathway. *PLoS Genet.* **6**, e1001250 doi:10.1371/journal.pgen.1001250

76  Ma, X.M., Yoon, S.-O., Richardson, C.J., Jülich, K. and Blenis, J. (2008) SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. *Cell* **133**, 303–313 doi:10.1016/j.cell.2008.02.031

77  Smyk, A., Szuminska, M., Uniewicz, K.A., Graves, L.M. and Kozlowski, P. (2006) Human enhancer of rudimentary is a molecular partner of PDIP46/SKAR, a protein interacting with DNA polymerase δ and S6K1 and regulating cell growth. *FEBS J.* **273**, 4728–4741 doi:10.1111/j.1742-4658.2006.05477.x

78  Orphanides, G., Wu, W.H., Lane, W.S., Hampsey, M. and Reinberg, D. (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* **400**, 284–288 doi:10.1038/22350

79  Bortvin, A. and Winston, F. (1996) Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* **272**, 1473–1476 doi:10.1126/science.272.5267.1473

80  Yoh, S.M., Cho, H., Pickle, L., Evans, R.M. and Jones, K.A. (2007) The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev.* **21**, 160–174 doi:10.1101/gad.1503107

81  Meinel, D.M., Burkert-Kautzsch, C., Kieser, A., O'Duibhir, E., Siebert, M., Mayer, A. et al. (2013) Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II. *PLoS Genet.* **9**, e1003914 doi:10.1371/journal.pgen.1003914

82  MacKellar, A.L. and Greenleaf, A.L. (2011) Cotranscriptional association of mRNA export factor Yra1 with C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **286**, 36385–36395 doi:10.1074/jbc.M111.268144

83  Gwizdek, C., Iglesias, N., Rodriguez, M.S., Ossareh-Nazari, B., Hobeika, M., Divita, G. et al. (2006) Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription. *Proc. Natl Acad. Sci. USA* **103**, 16376–16381 doi:10.1073/pnas.0607941103

84  Chanarat, S. and Sträßer, K. (2013) Splicing and beyond: the many faces of the Prp19 complex. *Biochim. Biophys. Acta* **1833**, 2126–2134 doi:10.1016/j.bbamcr.2013.05.023

85  Chanarat, S., Seizl, M. and Strasser, K. (2011) The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. *Genes Dev.* **25**, 1147–1158 doi:10.1101/gad.623411

86  David, C.J., Boyne, A.R., Millhouse, S.R. and Manley, J.L. (2011) The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev.* **25**, 972–983 doi:10.1101/gad.2038011

87  Dominguez-Sánchez, M.S., Barroso, S., Gómez-González, B., Luna, R. and Aguilera, A. (2011) Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet.* **7**, e1002386l doi:10.1371/journal.pgen.1002386

88  Davidson, L., Muniz, L. and West, S. (2014) 3′ end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes Dev.* **28**, 342–356 doi:10.1101/gad.231274.113

89  McCloskey, A., Taniguchi, I., Shinmyozu, K. and Ohno, M. (2012) hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science* **335**, 1643–1646 doi:10.1126/science.1218469

90  Luo, M.-J. and Reed, R. (1999) Splicing is required for rapid and efficient mRNA export in metazoans. *Proc. Natl Acad. Sci. USA* **96**, 14937–14942 doi:10.1073/pnas.96.26.14937

91  Takemura, R., Takeiwa, T., Taniguchi, I., McCloskey, A. and Ohno, M. (2011) Multiple factors in the early splicing complex are involved in the nuclear retention of pre-mRNAs in mammalian cells. *Genes Cells* doi:10.1111/j.1365-2443.2011.01548

92  Baejen, C., Torkler, P., Gressel, S., Essig, K., Söding, J. and Cramer, P. (2014) Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. *Mol. Cell.* **55**, 745–757 doi:10.1016/j.molcel.2014.08.005

93  Teng, I.-F. and Wilson, S.A. (2013) Mapping interactions between mRNA export factors in living cells. *PLoS ONE* **8**, e67676 doi:10.1371/journal.pone.0067676

94  Dias, A.P., Dufu, K., Lei, H. and Reed, R. (2010) A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat. Commun.* **1**, 97 doi:10.1038/ncomms1103

95  Le Hir, H., Izaurralde, E., Maquat, L.E. and Moore, M.J. (2000) The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**, 6860–6869 doi:10.1093/emboj/19.24.6860

96  Le Hir, H., Gatfield, D., Izaurralde, E. and Moore, M.J. (2001) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987–4997 doi:10.1093/emboj/20.17.4987

97  Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z. et al. (2012) The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151**, 750–764 doi:10.1016/j.cell.2012.10.007

98  Daguenet, E., Baguet, A., Degot, S., Schmidt, U., Alpy, F., Wendling, C. et al. (2012) Perispeckles are major assembly sites for the exon junction core complex. *Mol. Biol. Cell.* **23**, 1765–1782 doi:10.1091/mbc.E12-01-0040

99  Tuck, A.C. and Tollervey, D. (2013) A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. *Cell* **154**, 996–1009 doi:10.1016/j.cell.2013.07.047

100  Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K. et al. (2007) Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat. Chem. Biol.* **3**, 576–583 doi:10.1038/nchembio.2007.18

101  Martins, S.B., Rino, J., Carvalho, T., Carvalho, C., Yoshida, M., Klose, J.M. et al. (2011) Spliceosome assembly is coupled to RNA polymerase II dynamics at the 3′ end of human genes. *Nat. Struct. Mol. Biol.* **18**, 1115–1123 doi:10.1038/nsmb.2124

102  Nojima, T., Hirose, T., Kimura, H. and Hagiwara, M. (2007) The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. *J. Biol. Chem.* **282**, 15645–15651 doi:10.1074/jbc.M700629200

103  Blanchette, M., Labourier, E., Green, R.E., Brenner, S.E. and Rio, D.C. (2004) Genome-wide analysis reveals an unexpected function for the Drosophila splicing factor U2AF50 in the nuclear export of intronless mRNAs. *Mol. Cell.* **14**, 775–786 doi:10.1016/j.molcel.2004.06.012

104  Lei, H., Dias, A.P. and Reed, R. (2011) Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. *Proc. Natl Acad. Sci. USA* **108**, 17985–17990 doi:10.1073/pnas.1113076108

105  Lei, H., Zhai, B., Yin, S., Gygi, S. and Reed, R. (2013) Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export. *Nucleic Acids Res.* **41**, 2517–2525 doi:10.1093/nar/gks1314

106  Huang, Y. and Steitz, J.A. (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol. Cell.* **7**, 899–905 doi:10.1016/S1097-2765(01)00233-7

107  Hargous, Y., Hautbergue, G.M., Tintaru, A.M., Skrisovska, L., Golovanov, A.P., Stevenin, J. et al. (2006) Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *EMBO J.* **25**, 5126–5137 doi:10.1038/sj.emboj.7601385

108  Tintaru, A.M., Hautbergue, G.M., Hounslow, A.M., Hung, M.-L., Lian, L.-Y., Craven, C.J. et al. (2007) Structural and functional analysis of RNA and TAP binding to SF2/ASF. *EMBO Rep.* **8**, 756–762 doi:10.1038/sj.embor.7401031

109  Johnson, S.A., Kim, H., Erickson, B. and Bentley, D.L. (2011) The export factor Yra1 modulates mRNA 3′ end processing. *Nat. Struct. Mol. Biol.* **18**, 1164–1171 doi:10.1038/nsmb.2126

110  Rougemaille, M., Dieppois, G., Kisseleva-Romanova, E., Gudipati, R.K., Lemoine, S., Blugeon, C. et al. (2008) THO/sub2p functions to coordinate 3′-end processing with gene-nuclear pore association. *Cell* **135**, 308–321 doi:10.1016/j.cell.2008.08.005

111  Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, III, J.R. et al. (2009) Molecular architecture of the human pre-mRNA 3′ processing complex. *Mol. Cell.* **33**, 365–376 doi:10.1016/j.molcel.2008.12.028

112  Tran, D.D.H., Saran, S., Williamson, A.J.K., Pierce, A., Dittrich-Breiholz, O., Wiehlmann, L. et al. (2014) THOC5 controls 3′ end-processing of immediate early genes via interaction with polyadenylation specific factor 100 (CPSF100). *Nucleic Acids Res.* **42**, 12249–12260 doi:10.1093/nar/gku911

113  Katahira, J., Okuzaki, D., Inoue, H., Yoneda, Y., Maehara, K. and Ohkawa, Y. (2013) Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. *Nucleic Acids Res.* **41**, 7060–7072 doi:10.1093/nar/gkt414

PORTLAND
PRESS

114  Pak, V., Eifler, T.T., Jäger, S., Krogan, N.J., Fujinaga, K. and Peterlin, B.M. (2015) CDK11 in TREX/THOC regulates HIV mRNA 3′ end processing. *Cell Host Microbe* **18**, 560–570 doi:10.1016/j.chom.2015.10.012

115  Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J. et al. (2011) Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 doi:10.1038/nature10098

116  Gómez-González, B., García-Rubio, M., Bermejo, R., Gaillard, H., Shirahige, K., Marín, A. et al. (2011) Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *EMBO J.* **30**, 3106–3119 doi:10.1038/emboj.2011.206

117  Yoon, D.W., Lee, H., Seol, W., DeMaria, M., Rosenzweig, M. and Jung, J.U. (1997) Tap: a novel cellular protein that interacts with tip of Herpesvirus saimiri and induces lymphocyte aggregation. *Immunity* **6**, 571–582 doi:10.1016/S1074-7613(00)80345-3

118  Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Luhrmann, R. et al. (1997) Mex67p, a novel factor for nuclear mRNA export, binds to both poly (A)+ RNA and nuclear pores. *EMBO J.* **16**, 3256–3271 doi:10.1093/emboj/16.11.3256

119  Grüter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A. et al. (1998) TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell.* **1**, 649–659 doi:10.1016/S1097-2765(00)80065-9

120  Liker, E., Fernandez, E., Izaurralde, E. and Conti, E. (2000) The structure of the mRNA export factor TAP reveals a cis arrangement of a non-canonical RNP domain and an LRR domain. *EMBO J.* **19**, 5587–5598 doi:10.1093/emboj/19.21.5587

121  Aibara, S., Katahira, J., Valkov, E. and Stewart, M. (2015) The principal mRNA nuclear export factor NXF1:NXT1 forms a symmetric binding platform that facilitates export of retroviral CTE-RNA. *Nucleic Acids Res.* **43**, 1883–1893 doi:10.1093/nar/gkv032

122  Wickramasinghe, V.O., McMurtrie, P.I.A, Mills, A.D., Takei, Y., Penrhyn-Lowe, S., Amagase, Y. et al. (2010) mRNA export from mammalian cell nuclei is dependent on GANP. *Curr. Biol.* **20**, 25–31 doi:10.1016/j.cub.2009.10.078

123  Jani, D., Lutz, S., Hurt, E., Laskey, R.A., Stewart, M. and Wickramasinghe, V.O. (2012) Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic Acids Res.* **40**, 4562–4573 doi:10.1093/nar/gks059

124  Umlauf, D., Bonnet, J., Waharte, F., Fournier, M., Stierle, M., Fischer, B. et al. (2013) The human TREX-2 complex is stably associated with the nuclear pore basket. *J. Cell Sci.* **126**, 2656–2667 doi:10.1242/jcs.118000

125  Schneider, M., Hellerschmied, D., Schubert, T., Amlacher, S., Vinayachandran, V., Reja, R. et al. (2015) The nuclear pore-associated TREX-2 complex employs mediator to regulate gene expression. *Cell* **162**, 1016–1028 doi:10.1016/j.cell.2015.07.059

126  Rodríguez-Navarro, S., Fischer, T., Luo, M.-J., Antúnez, O., Brettschneider, S., Lechner, J. et al. (2004) Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116**, 75–86 doi:10.1016/S0092-8674(03)01025-0

127  Wickramasinghe, V.O., Andrews, R., Ellis, P., Langford, C., Gurdon, J.B., Stewart, M. et al. (2014) Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic Acids Res.* **42**, 5059–5071 doi:10.1093/nar/gku095

128  Wickramasinghe, V.O. and Laskey, R.A. (2015) Control of mammalian gene expression by selective mRNA export. *Nat. Rev. Mol. Cell Biol.* **16**, 431–442 doi:10.1038/nrm4010

129  Zolotukhin, A.S., Tan, W., Bear, J., Smulevitch, S. and Felber, B.K. (2002) U2AF participates in the binding of TAP (NXF1) to mRNA. *J. Biol. Chem.* **277**, 3935–3942 doi:10.1074/jbc.M107598200

130  Katahira, J., Dimitrova, L., Imai, Y. and Hurt, E. (2015) NTF2-like domain of Tap plays a critical role in cargo mRNA recognition and export. *Nucleic Acids Res.* **43**, 1894–1904 doi:10.1093/nar/gkv039

131  Li, Y., Bor, Y.-C., Misawa, Y., Xue, Y., Rekosh, D. and Hammarskjöld, M.-L. (2006) An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature* **443**, 234–237 doi:10.1038/nature05107

132  Huang, Y., Gattoni, R., Stévenin, J. and Steitz, J.A. (2003) SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell.* **11**, 837–843 doi:10.1016/S1097-2765(03)00089-3

133  Hurt, E., Luo, M.-J., Röther, S., Reed, R. and Sträßer, K. (2004) Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proc. Natl Acad. Sci. USA* **101**, 1858–1862 doi:10.1073/pnas.0308663100

134  Reed, R. and Cheng, H. (2005) TREX, SR proteins and export of mRNA. *Curr. Opin. Cell Biol.* **17**, 269–273 doi:10.1016/j.ceb.2005.04.011

135  Martínez-Lumbreras, S., Taverniti, V., Zorrilla, S., Séraphin, B. and Pérez-Cañadillas, J.M. (2016) Gbp2 interacts with THO/TREX through a novel type of RRM domain. *Nucleic Acids Res.* **44**, 437–448 doi:10.1093/nar/gkv1303

136  Lindtner, S., Zolotukhin, A.S., Uranishi, H., Bear, J., Kulkarni, V., Smulevitch, S. et al. (2006) RNA-binding motif protein 15 binds to the RNA transport element RTE and provides a direct link to the NXF1 export pathway. *J. Biol. Chem.* **281**, 36915–36928 doi:10.1074/jbc.M608745200

137  Ruepp, M.-D., Aringhieri, C., Vivarelli, S., Cardinale, S., Paro, S., Schumperli, D. et al. (2009) Mammalian pre-mRNA 3′ end processing factor CF Im 68 functions in mRNA export. *Mol. Biol. Cell.* **20**, 5211–5223 doi:10.1091/mbc.E09-05-0389

138  Qu, X., Lykke-Andersen, S., Nasser, T., Saguez, C., Bertrand, E., Jensen, T.H. et al. (2009) Assembly of an export-competent mRNP is needed for efficient release of the 3′-end processing complex after polyadenylation. *Mol. Cell Biol.* **29**, 5327–5338 doi:10.1128/MCB.00468-09

139  Hurt, J.A., Obar, R.A., Zhai, B., Farny, N.G., Gygi, S.P. and Silver, P.A. (2009) A conserved CCCH-type zinc finger protein regulates mRNA nuclear adenylation and export. *J. Cell Biol.* **185**, 265–277 doi:10.1083/jcb.200811072

140  Folkmann, A.W., Noble, K.N., Cole, C.N. and Wente, S.R. (2011) Dbp5, Gle1-IP6, and Nup159: a working model for mRNP export. *Nucleus* **2**, 540–548 doi:10.4161/nucl.2.6.17881

141  Schmitt, C., von Kobbe, C., Bachi, A., Pante, N., Rodrigues, J.P., Boscheron, C. et al. (1999) Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. *EMBO J.* **18**, 4332–4347 doi:10.1093/emboj/18.15.4332

142  Folkmann, A.W., Collier, S.E., Zhan, X., Aditi, X., Ohi, M.D. and Wente, S.R. (2013) Gle1 functions during mRNA export in an oligomeric complex that is altered in human disease. *Cell* **155**, 582–593 doi:10.1016/j.cell.2013.09.023

143  Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B.S., Schultz, N. et al. (2012) UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151**, 871–884 doi:10.1016/j.cell.2012.09.040

144  Hur, J.K., Luo, Y., Moon, S., Ninova, M., Marinov, G.K., Chung, Y.D. et al. (2016) Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in Drosophila. *Genes Dev.* **30**, 840–855 doi:10.1101/gad.276030.115

145  Santos-Pereira, J.M. and Aguilera, A. (2015) R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* **16**, 583–597 doi:10.1038/nrg3961

146 Wickramasinghe, V.O. and Venkitaraman, A.R. (2016) RNA processing and genome stability: cause and consequence. *Mol. Cell.* **61**, 496–505 doi:10.1016/j.molcel.2016.02.001

147 Huertas, P. and Aguilera, A. (2003) Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell.* **12**, 711–721 doi:10.1016/j.molcel.2003.08.010

148 Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A. et al. (2009) A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol. Cell.* **35**, 228–239 doi:10.1016/j.molcel.2009.06.021

149 Nedea, E., Nalbant, D., Xia, D., Theoharis, N.T., Suter, B., Richardson, C.J. et al. (2008) The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes. *Mol. Cell.* **29**, 577–587 doi:10.1016/j.molcel.2007.12.031

150 Wickramasinghe, V.O., Savill, J.M., Chavali, S., Jonsdottir, A.B., Rajendra, E., Grüner, T. et al. (2013) Human inositol polyphosphate multikinase regulates transcript-selective nuclear mRNA export to preserve genome integrity. *Mol. Cell.* **51**, 737–750 doi:10.1016/j.molcel.2013.08.031

151 Takai, H., Masuda, K., Sato, T., Sakaguchi, Y., Suzuki, T., Suzuki, T. et al. (2014) 5-Hydroxymethylcytosine plays a critical role in glioblastomagenesis by recruiting the CHTOP-methylosome complex. *Cell Rep.* **9**, 48–60 doi:10.1016/j.celrep.2014.08.071

152 Li, Y., Lin, A.W., Zhang, X., Wang, Y., Wang, X. and Goodrich, D.W. (2007) Cancer cells and normal cells differ in their requirements for Thoc1. *Cancer Res.* **67**, 6657–6664 doi:10.1158/0008-5472.CAN-06-3234

153 Guo, S., Hakimi, M.A., Baillat, D., Chen, X., Farber, M.J., Klein-Szanto, A.J. et al. (2005) Linking transcriptional elongation and messenger RNA export to metastatic breast cancers. *Cancer Res.* **65**, 3011–3016 PMID: 15833825

154 Domínguez-Sánchez, M.S., Sáez, C., Japón, M.A., Aguilera, A. and Luna, R. (2011) Differential expression of THOC1 and ALY mRNP biogenesis/export factors in human cancers. *BMC Cancer* **11**, 77 doi:10.1186/1471-2407-11-77

155 Guo, S., Liu, M. and Godwin, A.K. (2012) Transcriptional regulation of hTREX84 in human cancer cells. *PLoS ONE* **7**, e43610 doi:10.1371/journal.pone.0043610

156 Griaud, F., Pierce, A., Gonzalez Sanchez, M.B., Scott, M., Abraham, S.A., Holyoake, T.L. et al. (2013) A pathway from leukemogenic oncogenes and stem cell chemokines to RNA processing via THOC5. *Leukemia* **27**, 932–940 doi:10.1038/leu.2012.283

157 Saito, Y., Kasamatsu, A., Yamamoto, A., Shimizu, T., Yokoe, H., Sakamoto, Y. et al. (2013) ALY as a potential contributor to metastasis in human oral squamous cell carcinoma. *J. Cancer Res. Clin. Oncol.* **139**, 585–594 doi:10.1007/s00432-012-1361-5

158 Crawford, N.P.S., Qian, X., Ziogas, A., Papageorge, A.G., Boersma, B.J., Walker, R.C. et al. (2007) Rrp1b, a new candidate susceptibility gene for breast cancer progression and metastasis. *PLoS Genet.* **3**, e214 doi:10.1371/journal.pgen.0030214

159 Crawford, N.P., Yang, H., Mattaini, K.R. and Hunter, K.W. (2009) The metastasis efficiency modifier ribosomal RNA processing 1 homolog B (RRP1B) is a chromatin-associated factor. *J. Biol. Chem.* **284**, 28660–28673 doi:10.1074/jbc.M109.023457

160 Tsai, Y.C. and Weissman, A.M. (2011) Dissecting the diverse functions of the metastasis suppressor CD82/KAI1. *FEBS Lett.* **585**, 3166–3173 doi:10.1016/j.febslet.2011.08.031

161 Choong, M.L., Tan, L.K., Lo, S.L., Ren, E.-C., Ou, K., Ong, S.-E. et al. (2001) An integrated approach in the discovery and characterization of a novel nuclear protein over-expressed in liver and pancreatic tumors. *FEBS Lett.* **496**, 109–116 doi:10.1016/S0014-5793(01)02409-7

162 Hashii, Y., Kim, J.Y., Sawada, A., Tokimasa, S., Hiroyuki, F., Ohta, H. et al. (2004) A novel partner gene CIP29 containing a SAP domain with MLL identified in infantile myelomonocytic leukemia. *Leukemia* **18**, 1546–1548 doi:10.1038/sj.leu.2403450

163 Li, B.E. and Ernst, P. (2014) Two decades of leukemia oncoprotein epistasis: the MLL1 paradigm for epigenetic deregulation in leukemia. *Exp. Hematol.* **42**, 995–1012 doi:10.1016/j.exphem.2014.09.006

164 Osinalde, N., Olea, M., Mitxelena, J., Aloria, K., Rodriguez, J.A., Fullaondo, A. et al. (2013) The nuclear protein ALY binds to and modulates the activity of transcription factor E2F2. *Mol. Cell. Proteomics* **12**, 1087–1098 doi:10.1074/mcp.M112.024158

165 Goshima, G., Wollman, R., Goodwin, S.S., Zhang, N., Scholey, J.M., Vale, R.D. et al. (2007) Genes required for mitotic spindle assembly in Drosophila S2 cells. *Science* **316**, 417–421 doi:10.1126/science.1141314

166 Somma, M.P., Ceprani, F., Bucciarelli, E., Naim, V., De Arcangelis, V., Piergentili, R. et al. (2008) Identification of Drosophila mitotic genes by combining co-expression analysis and RNA interference. *PLoS Genet.* **4**, e1000126 doi:10.1371/journal.pgen.1000126

167 Wang, X., Chang, Y., Li, Y., Zhang, X. and Goodrich, D.W. (2006) Thoc1/Hpr1/p84 is essential for early embryonic development in the mouse. *Mol. Cell. Biol.* **26**, 4362–4367 doi:10.1128/MCB.02163-05

168 Pitzonka, L., Wang, X., Ullas, S., Wolff, D.W., Wang, Y. and Goodrich, D.W. (2013) The THO ribonucleoprotein complex is required for stem cell homeostasis in the adult mouse small intestine. *Mol. Cell. Biol.* **33**, 3505–3514 doi:10.1128/MCB.00751-13

169 Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H. and Elledge, S.J. (2009) A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev.* **23**, 837–848 doi:10.1101/gad.1769609

170 Carney, L., Pierce, A., Rijnen, M., Gonzalez Sanchez, M.B., Hamzah, H.G., Zhang, L. et al. (2009) THOC5 couples M-CSF receptor signaling to transcription factor expression. *Cell Signal.* **21**, 309–316 doi:10.1016/j.cellsig.2008.10.018

171 Tran, D.D.H., Saran, S., Dittrich-Breiholz, O., Williamson, A.J.K., Klebba-Färber, S., Koch, A. et al. (2013) Transcriptional regulation of immediate-early gene response by THOC5, a member of mRNA export complex, contributes to the M-CSF-induced macrophage differentiation. *Cell Death Dis.* **4**, e879 doi:10.1038/cddis.2013.409

172 Saran, S., Tran, D.D.H., Klebba-Färber, S., Moran-Losada, P., Wiehlmann, L., Koch, A. et al. (2013) THOC5, a member of the mRNA export complex, contributes to processing of a subset of wingless/integrated (Wnt) target mRNAs and integrity of the gut epithelial barrier. *BMC Cell Biol.* **14**, 51 doi:10.1186/1471-2121-14-51

173 Wang, X., Chinnam, M., Wang, J., Wang, Y., Zhang, X., Marcon, E. et al. (2009) Thoc1 deficiency compromises gene expression necessary for normal testis development in the mouse. *Mol. Cell. Biol.* **29**, 2794–2803 doi:10.1128/MCB.01633-08

174 Beaulieu, C.L., Huang, L., Innes, A, Akimenko, M.-A., Puffenberger, E.G., Schwartz, C. et al. (2013) Intellectual disability associated with a homozygous missense mutation in THOC6. *Orphanet J. Rare Dis.* **8**, 62 doi:10.1186/1750-1172-8-62

175 Boycott, K.M., Beaulieu, C., Puffenberger, E.G. McLeod, D.R., Parboosingh, J.S. and Innes, A.M. (2010) A novel autosomal recessive malformation syndrome associated with developmental delay and distinctive facies maps to 16ptel in the Hutterite population. *Am. J. Med. Genet.* **152A**, 1349–1356 doi:10.1002/ajmg.a.33379

PORTLAND
PRESS

176 Walsh, M.J., Cooper-Knock, J., Dodd, J.E., Stopford, M.J., Mihaylov, S.R., Kirby, J. et al. (2015) Invited review: decoding the pathophysiological mechanisms that underlie RNA dysregulation in neurodegenerative disorders: a review of the current state of the art. *Neuropathol. Appl. Neurobiol.* **41**, 109–134 doi:10.1111/nan.12187

177 Rohrer, J.D., Isaacs, A.M., Mizielinska, S., Mead, S., Lashley, T., Wray, S. et al. (2015) C9orf72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. *Lancet Neurol.* **14**, 291–301 doi:10.1016/S1474-4422(14)70233-9

178 Cooper-Knock, J., Walsh, M.J., Higginbottom, A., Robin Highley, J., Dickman, M.J., Edbauer, D. et al. (2014) Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* **137**, 2040–2051 doi:10.1093/brain/awu120

179 Freibaum, B.D., Lu, Y., Lopez-Gonzalez, R., Kim, N.C., Almeida, S., Lee, K.H. et al. (2015) GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* doi:10.1038/nature14974

180 Read, E.K. and Digard, P. (2010) Individual influenza A virus mRNAs show differential dependence on cellular NXF1/TAP for their nuclear export. *J. Gen. Virol.* **91**, 1290–1301 doi:10.1099/vir.0.018564-0

181 Larsen, S., Bui, S., Perez, V., Mohammad, A., Medina-Ramirez, H. and Newcomb, L.L. (2014) Influenza polymerase encoding mRNAs utilize atypical mRNA nuclear export. *Virol. J.* **11**, 154 doi:10.1186/1743-422X-11-154

182 Malik, P., Blackbourn, D.J. and Clements, J.B. (2004) The evolutionarily conserved Kaposi's sarcoma-associated Herpesvirus ORF57 protein interacts with REF protein and acts as an RNA export factor. *J. Biol. Chem.* **279**, 33001–33011 doi:10.1074/jbc.M313008200

183 Williams, B.J.L, Boyne, J.R., Goodwin, D.J., Roaden, L., Hautbergue, G.M., Wilson, S.A. et al. (2005) The prototype gamma-2 herpesvirus nucleocytoplasmic shuttling protein, ORF 57, transports viral RNA via the cellular mRNA export pathway. *Biochem. J.* **387**, 295–308 doi:10.1042/BJ20041223

184 Boyne, J.R., Colgan, K.J. and Whitehouse, A. (2008) Recruitment of the complete hTREX complex is required for Kaposi's sarcoma-associated Herpesvirus intronless mRNA nuclear export and virus replication. *PLoS Pathog.* **4**, e1000194 doi:10.1371/journal.ppat.1000194

185 Tunnicliffe, R.B., Hautbergue, G.M., Kalra, P., Jackson, B.R., Whitehouse, A., Wilson, S.A. et al. (2011) Structural basis for the recognition of cellular mRNA export factor REF by herpes viral proteins HSV-1 ICP27 and HVS ORF57. *PLoS Pathog.* **7**, e1001244 doi:10.1371/journal.ppat.1001244

186 Jackson, B.R., Boyne, J.R., Noerenberg, M., Taylor, A., Hautbergue, G.M., Walsh, M.J. et al. (2011) An interaction between KSHV ORF57 and UIF provides mRNA-adaptor redundancy in Herpesvirus intronless mRNA export. *PLoS Pathog.* **7**, e1002138 doi:10.1371/journal.ppat.1002138

187 Koffa, M.D., Clements, J.B., Izaurralde, E., Wadd, S., Wilson, S.A., Mattaj, I.W. et al. (2001) Herpes simplex virus ICP27 protein provides viral mRNAs with access to the cellular mRNA export pathway. *EMBO J.* **20**, 5769–5778 doi:10.1093/emboj/20.20.5769

188 Hiriart, E., Farjot, G., Gruffat, H., Nguyen, M.V., Sergeant, A. and Manet, E. (2002) A novel nuclear export signal and a REF interaction domain both promote mRNA export by the Epstein-Barr virus EB2 protein. *J. Biol. Chem.* **278**, 335–342 doi:10.1074/jbc.M208656200

189 Lischka, P., Toth, Z., Thomas, M., Mueller, R. and Stamminger, T. (2006) The UL69 transactivator protein of human cytomegalovirus interacts with DEXD/H-Box RNA helicase UAP56 to promote cytoplasmic accumulation of unspliced RNA. *Mol. Cell. Biol.* **26**, 1631–1643 doi:10.1128/MCB.26.5.1631-1643.2006

190 Balasubramaniam, V.R., Hong Wai, T., Ario Tejo, B., Omar, A.R. and Syed Hassan, S. (2013) Highly pathogenic avian influenza virus nucleoprotein interacts with TREX complex adaptor protein Aly/REF. *PLoS ONE* **8**, e72429 doi:10.1371/journal.pone.0072429

191 Stubbs, S.H., Hunter, O.V., Hoover, A. and Conrad, N.K. (2012) Viral factors reveal a role for REF/Aly in nuclear RNA stability. *Mol. Cell Biol.* **32**, 1260–1270 doi:10.1128/MCB.06420-11

192 Kammel, C., Thomaier, M., Sørensen, B.B., Schubert, T., Längst, G., Grasser, M. et al. (2013) Arabidopsis DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. *PLoS ONE* **8**, e60644 doi:10.1371/journal.pone.0060644

193 Pan, H., Liu, S. and Tang, D. (2012) HPR1, a component of the THO/TREX complex, plays an important role in disease resistance and senescence in Arabidopsis. *Plant J.* **69**, 831–843 doi:10.1111/j.1365-313X.2011.04835.x

194 Jauvion, V., Elmayan, T. and Vaucheret, H. (2010) The conserved RNA trafficking proteins HPR1 and TEX1 are involved in the production of endogenous and exogenous small interfering RNA in Arabidopsis. *Plant Cell.* **22**, 2697–2709 doi:10.1105/tpc.110.076638

195 Xu, C., Zhou, X. and Wen, C.K. (2015) HYPER RECOMBINATION1 of the THO/TREX complex plays a role in controlling transcription of the REVERSION-TO-ETHYLENE SENSITIVITY1 gene in Arabidopsis. *PLoS Genet.* **11**, e1004956 doi:10.1371/journal.pgen.1004956

196 Pitzonka, L., Wang, X., Ullas, S., Wolff, D.W., Wang, Y. and Goodrich, D.W. (2013) The THO ribonucleoprotein complex is required for stem cell homeostasis in the adult mouse small intestine. *Mol. Cell. Biol.* **33**, 3505–3514

197 Isomura, M., Okui, K., Fujiwara, T., Shin, S. and Nakamura, Y. (1996) Cloning and mapping of a novel human cDNA homologous to DROER, the enhancer of the Drosophila melanogaster rudimentary gene. *Genomics* **32**, 125–127

198 Black, B.E., Levesque, L., Holaska, J.M., Wood, T.C. and Paschal, B.M. (1999) Identification of an NTF2-related factor that binds Ran-GTP and regulates nuclear protein export. *Mol. Cell. Biol.* **19**, 8616–8624

199 Guzik, B.W., Levesque, L., Prasad, S., Bor, Y.C., Black, B.E., Paschal, B.M. et al. (2001) NXT1 (p15) is a crucial cellular cofactor in TAP-dependent export of intron-containing RNA in mammalian cells. *Mol. Cell. Biol.* **21**, 2545–2554

200 Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S. et al. (2011) Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell.* **43**, 624–637

201 Wilson, M.D., Wang, D., Wagner, R., Breyssens, H., Gertsenstein, M., Lobe, C. et al. (2008) ARS2 is a conserved eukaryotic gene essential for early mammalian development. *Mol. Cell. Biol.* **28**, 1503–1514

202 Gruber, J.J., Zatechka, D.S., Sabin, L.R., Yong, J., Lum, J.J., Kong, M. et al. (2009) Ars2 links the nuclear cap-binding complex to RNA interference and cell proliferation. *Cell* **138**, 328–339

203 Hallais, M., Pontvianne, F., Andersen, P.R., Clerici, M., Lener, D., Benbahouche, N.H. et al. (2013) CBC-ARS2 stimulates 3'-end maturation of multiple RNA families and favors cap-proximal processing. *Nat. Struct. Mol. Biol.* **20**, 1358–1366

204 Merz, C., Urlaub, H., Will, C.L. and Lührmann, R. (2007) Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *RNA* **13**, 116–128

205 Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E. and Mattaj, I.W. (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78**, 657–668