BIOINFORMATICS ORIGINAL PAPER

Vol. 31 no. 2 2015, pages 246–251
doi:10.1093/bioinformatics/btu620

Databases and ontologies

Advance Access publication September 30, 2014

ExportAid: database of RNA elements regulating nuclear RNA export in mammals

Matteo Giulietti, Sara Armida Milantoni, Tatiana Armeni, Giovanni Principato and Francesco Piva*

Department of Specialistic Clinical and Odontostomatological Sciences, Polytechnic University of Marche, Ancona 60131, Italy

Associate Editor: Jonathan Wren

ABSTRACT

Motivation: Regulation of nuclear mRNA export or retention is carried out by RNA elements but the mechanism is not yet well understood. To understand the mRNA export process, it is important to collect all the involved RNA elements and their trans-acting factors.

Results: By hand-curated literature screening we collected, in ExportAid database, experimentally assessed data about RNA elements regulating nuclear export or retention of endogenous, heterologous or artificial RNAs in mammalian cells. This database could help to understand the RNA export language and to study the possible export efficiency alterations owing to mutations or polymorphisms. Currently, ExportAid stores 235 and 96 RNA elements, respectively, increasing and decreasing export efficiency, and 98 neutral assessed sequences.

Availability and implementation: Freely accessible without registration at http://www.intron i.it/ExportAid/ExportAid.html. Database and web interface are implemented in Perl, MySQL, Apache and JavaScript with all major browsers supported.

Contact: f.piva@univpm.it

Received on April 14, 2014; revised on July 30, 2014; accepted on September 14, 2014

1 INTRODUCTION

The major challenge of the post-genomic era is to understand the meaning of the genomic sequences, i.e. to decipher the languages used by different molecular processes such as transcription, RNA and protein folding, RNA splicing, RNA export, microRNA regulation and so on. Unfortunately, this research is demanding because each language has a wide set of possible words, they are not placed in fixed positions in genes, the words of different languages could overlap and many words have more than one meaning because of the cellular context. For example, genetic, RNA splicing, RNA folding and RNA export codes can coexist in exons (Barbieri, 2003; Giulietti *et al.*, 2013; Piva *et al.*, 2012).

A highly conserved process among eukaryotes but not yet well understood is the RNA export. It occurs at the end of transcripts' maturation, finely regulating their nuclear export, retention or import (Carmody and Wente, 2009; Cullen, 2003; Erkmann and Kutay, 2004). Small RNAs are bound directly by the mobile export receptors that bring RNAs through the nuclear pore complex, whereas mRNAs assemble into ribonucleoprotein particles and recruit their exporter receptors, such as TAP–p15 and CRM1, via adaptor proteins like ALY/REF and SR proteins (Anko, 2014; Kohler and Hurt, 2007; Stewart, 2010). Most of the adaptors are part of the exon junction complex (EJC), which, binding non-specific sequences near exonic junctions, mediates the mRNA export (Le Hir *et al.*, 2001; Rodrigues *et al.*, 2001; Zhou *et al.*, 2000). However, the EJC role is not completely clear; in fact, one could wonder if transcripts with many introns are exported in a more efficient manner, as, owing to splicing, they manage to recruit more EJC proteins. Alternatively, more EJC proteins would be needed to carry out the export of long mRNAs.

As also transcripts that do not undergo splicing can be exported, this suggests that further *cis*-acting RNA elements provide for the mRNA export. Actually, by studying mammalian intronless RNAs (Lei *et al.*, 2011, 2013) and some viral transcripts, mainly in retroviruses (Leblanc *et al.*, 2013), new RNA export elements have been discovered.

It emerges that a comprehensive knowledge of all the RNA sequences involved in nuclear export, and of their *trans*-acting factors, is crucial for a better understanding of the transcript's nuclear export process and functions. In particular, collecting all this information, scattered throughout the literature, in a unique resource would help to explain the export of transcripts, such as intronless genes, some long non-coding RNAs (lncRNAs) and viral genomes that do not undergo splicing, or the retention of transcripts that undergo splicing. Moreover, to shed light on this mechanism, it would be useful to study the possible export efficiency alterations owing to mutations or polymorphisms.

By hand-curated literature screening, we collected, in ExportAid database, experimentally assessed data about RNA elements regulating nuclear export or retention of endogenous, heterologous or artificial RNAs in mammalian cells. This specialized resource constitutes a repository, including information on functional sequences, their genomic location, experimental data and involved export pathways. ExportAid is freely accessible at www.intron i.it/ExportAid/ExportAid.html and currently stores 235 and 96 RNA elements, respectively, increasing and decreasing export efficiency, and 98 neutral assessed sequences.

*To whom correspondence should be addressed.

2 SYSTEM AND METHODS

For PubMed searches, we entered the words 'RNA export', 'export RNA element' or other similar terms or the name of the known RNA export elements, including their synonyms. We included all papers reporting mammalian, viral, mutated or artificial RNA elements reported to be functional for the RNA export in mammalian cell lines. For each paper, we annotated the reported sequences and related information, such as the name of the RNA functional elements lying in the sequences, the organism to which the sequences belong and the genomic coordinates along with the GenBank accession numbers. We also extracted the experimental conditions, such as the cell lines used to test the sequence functionality, and the structure of the genetic constructs used in transfections. In particular, we annotated whether the constructs were produced by joining sequences from different parts of the same genome or from different genomes, whether constructs were built by putting together different functional elements and whether the sequences belonged to known genes. As constructs are usually cited in papers by name, to be easily recognized, we also collected their name, or when missing, the name of the used vector.

As in all papers the authors quantified the export efficiency of each RNA element, by the percentage of relative cytoplasmic accumulation by assessing the RNA or the activity level of a reporter gene, we thought it would have been useful to annotate these functionality levels. Therefore, for each paper, we scaled them from 1 to 10 and assigned a score ranging from –10 to +10 for each sequence. This score describes the functional effect of an RNA element in terms of tendency to increase or decrease the nuclear export efficiency. We annotated '(0) neutral effect' to those sequences for which insertion or deletion in a construct did not alter the export efficiency. We assigned a positive or a negative score to the functional elements that, respectively, increased or decreased the RNA export efficiency. The user has to consider that an RNA element with score +10 does not confer 100% export efficiency of the construct, but only that it has the maximum effect in increasing the export among the sequences assessed in the same paper. The same goes for a score of –10.

Finally, to assure the database quality, we established guidelines for the literature review and annotation of records, and performed an inter-annotator analysis (Colosimo *et al.*, 2005) carried out by post-doctoral researchers, PhD students and a senior researcher.

3 DATABASE

3.1 Current status of the database

Overall, the database contains 235 and 96 RNA sequences, respectively, increasing and decreasing export efficiency, and 98 neutral assessed sequences. As shown in Table 1, most of the elements derive from viral genomes, in particular HIV and SRV retrovirus, whereas humans have a bigger variety of functional elements. It should be noted that we have also included information about the post-transcriptional regulatory element of woodchuck hepatitis virus (WPRE), although its involvement in

Table 1. Distribution of the RNA elements collected in ExportAid

Organism	RNA element	Count
Human	CAR (cytoplasmic accumulation region)	10
	CJE (c-jun gene enhancer)	3
	CTE	7
	eIF4E-SE (eIF4E sensitive element)	3
	PRE	1
	SSCR	2
	Part of genome/synthetic sequence	1
Mouse	H2a element	5
	MTE (musD transport element)	8
	RTE (RNA transport eelement)	24
	SSCR	3
	Part of genome	3
ALV	CTE	5
CAEV	RRE-C	1
	part of genome/synthetic sequence	2
EIAV	RRE	10
FIV	RRE	1
HBV	PRE or HPRE	42
	HBV element	3
	Export promoting domain β	10
	PPD1 - polypyrimidine domain 1	9
	PPD2 - polypyrimidine domain 2	9
	Part of genome/synthetic sequence	3
HERV-K (HTDV)	K-RRE	2
HHV-8 (KSHV)	KSHV element	2
HIV-1	INS (inhibitory/instability sequence, CRS or IR)	77
	RRE	65
	Part of genome/synthetic sequence	16
HIV-2	INS (inhibitory/instability sequence, CRS or IR)	1
	Part of genome/synthetic sequence	6
HPV	Part of genome/synthetic sequence	16
HSV-1	PPE (pre-mRNA processing enhancer)	15
HTLV-1	CIE (*cis*-acting inhibitory element)	21
	RxRE (Rex-responsive element, also XRE)	7
	Part of genome/synthetic sequence	7
MPMV	CTE	9
RSV	DR (direct repeat)	24
	Part of genome	1
SRV-1	CTE	44
SRV-2	CTE	2
WHV	WPRE (WHV PRE)	2
Synthetic sequences		6

Notes: 'part of genome/synthetic sequence' indicates that the assessed sequence did not harbour known functional elements. ALV, avian leukosis virus; CAEV, caprine arthritis-encephalitis virus; EIAV, equine infectious anaemia virus; FIV, feline immunodeficiency virus; HBV, hepatitis B virus; HERV-K, human endogenous retrovirus type K, also HTDV; HHV-8/KSHV, human herpesvirus type 8 / Kaposi's sarcoma-associated herpesvirus; HIV-1, (human immunodeficiency virus type 1; HIV-2, human immunodeficiency virus type 2; HPV, human papillomavirus types 1, 2, 6b, 16, 18, 31, 41 and 61; HSV-1, herpes simplex virus type 1; HTLV-1, human T-lymphotropic virus type 1; MPMV, Mason–Pfizer monkey virus; RSV, Rous sarcoma virus; SRV-1, simian retrovirus 1; SRV-2, simian retrovirus 2; WHV, woodchuck hepatitis virus.

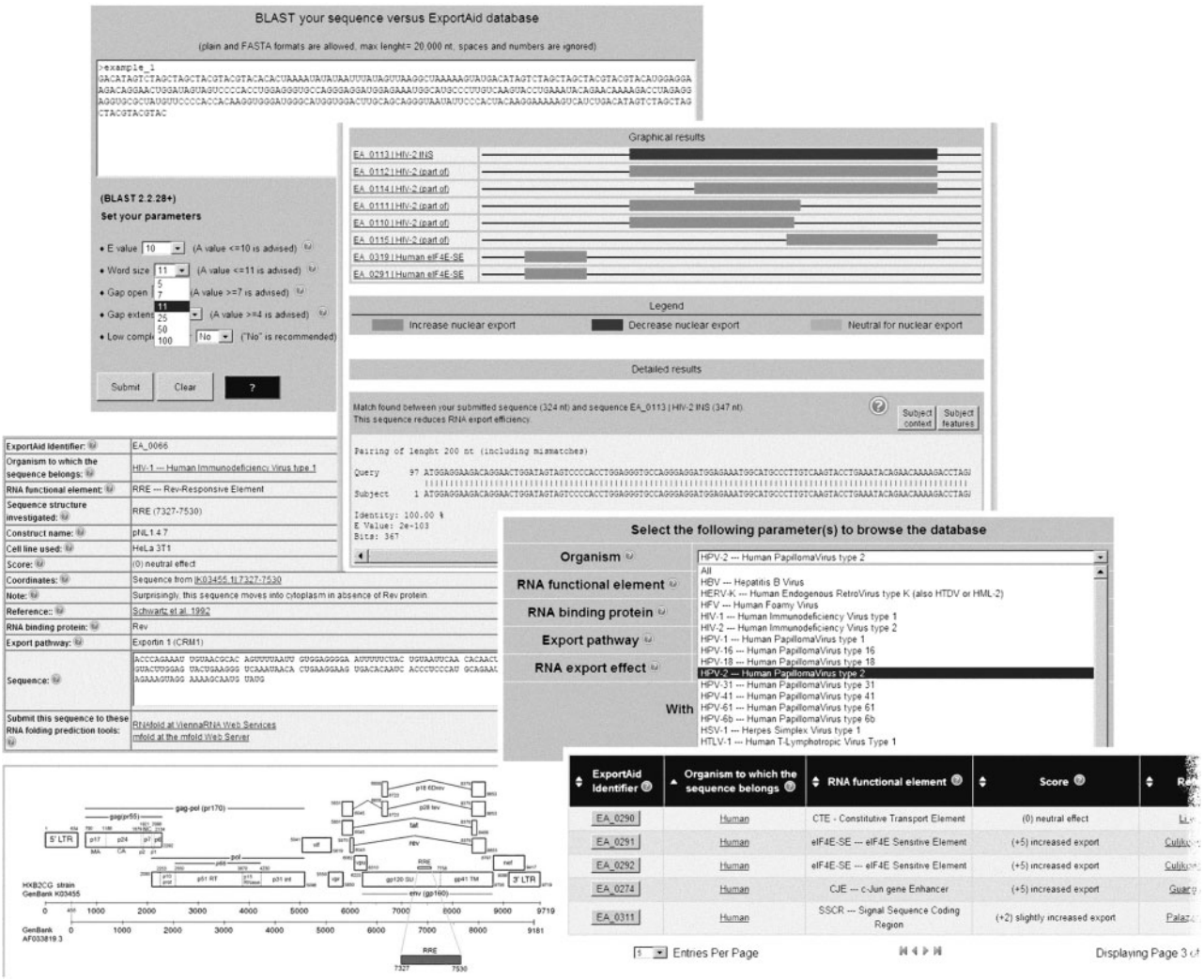

Fig. 1. ExportAid web interfaces. At the top: Blast page and resulting graphical alignments. At the bottom: Browse page and resulting records and cards

export is still debated (Donello *et al.*, 1998; Higashimoto *et al.*, 2007; Zufferey *et al.*, 1999).

Unfortunately for most functional elements, there are few known sequences. Even the PRE has been found, to date, only in CD83 human gene (Prechtel *et al.*, 2006). Instead, a few export elements have been found in various genes that, intriguingly, are involved in similar functions. For example, the sequence element eIF4E-SE (eIF4E sensitive element) is borne by some genes, mainly those encoding cyclins, that constitute an RNA regulon governing cellular proliferation (Culjkovic *et al.*, 2006). Analogously, the signal sequence coding region (SSCR) element lies in mRNAs coding for secretory proteins (Palazzo *et al.*, 2007).

Most RNA export elements we collected have a length ranging from a few ten to a few hundred nucleotides. The shortest ExportAid sequences are mouse H2a element (22 nt) and human eIF4E sensitive element (40 nt). Among the characteristic RNA export elements used by retroviruses to express their unspliced mRNAs in the host cell, the constitutive transport element (CTE) of the simian retrovirus and Mason–Pfizer monkey virus and the RRE (Rev-responsive element) of the human immunodeficiency virus type 1 are the most studied. Moreover, we annotated whether some retroviruses encode the *trans*-acting factors that interact with the virus's export elements, such as Rev protein of the human immunodeficiency virus or ORF57 of the human herpesvirus-8.

3.2 Using ExportAid

ExportAid can be freely queried using Browse or Blast pages. The first page contains different criteria for performing searches (by organism, RNA element, RNA-binding protein, export pathway and functional effect), and at the bottom of the page, it is possible to download the entire database in different formats. Results are returned as concise records, each corresponding to an RNA export sequence, and the user can expand the record of interest, thus obtaining the card with the detailed annotations (Fig. 1).

We have installed the local NCBI BLAST, which is accessible via the Blast page, to allow the user to search his sequence versus

the entire database returning alignments linked to the respective cards. The returned alignments have to be correctly interpreted because not all alignments mean the presence of an RNA functional element in the user sequence. It should be taken into account that ExportAid collects the exact sequences assessed in the reference papers, so some of them consist of hundreds or thousands of nucleotides, but only a shorter part is thought to be functional. To help the user to identify the functional elements inside a sequence, we equipped each card with an image we made that shows the structure and the genomic context each sequence derives from. So the user can verify whether the returned alignment overlaps to a known functional element by using the images enclosed with each card. The colour of the boxes representing functional elements reflects their main activity: for example, the INS element of HIV-1 is red because usually it prevents export, and the RRE element of HIV-1 is green because generally it promotes export. In some cases, an RNA export element can exert opposite functions depending on the presence or absence of its *trans*-acting factor, and the user can retrieve this information from the 'Note' field. For example, the RRE element of HIV-1 has a context-dependent functionality, as it prevents or enhances nuclear export respectively in absence or in presence of its Rev protein.

Naturally, if the returned alignment does not extend to the entire length of the functional element, it does not guarantee the functionality of the user-submitted sequence, as the binding site for its *trans*-acting factor could be lacking or its binding could be prevented. In fact, RNA secondary structure is important for the function, and a shorter sequence would not fold correctly. To make these evaluations easier, we added some links in each card that allow the user to fold the database sequence by different RNA folding tools, such as RNAfold, mfold, UNAFold, RNAshapes, Context Fold and RNAstructure.

Regarding annotations of the *trans*-acting elements, we took into account also the following rare cases. When a sequence did not contain any known functional element and was neutral for export according to the literature, we indicated 'None' in the fields 'RNA functional element' and 'RNA binding protein', as we assumed that no RNA binding protein involved in export could bind this sequence. When an RNA element lying in a sequence was unknown but the sequence was assessed as functional, we indicated 'Unknown' in the aforementioned fields, as we assumed that the sequence could bind some *trans*-acting factor involved in export but was not yet identified.

Annotating genomic coordinates, we realized that some viral sequences are associated to different genome assemblies, and therefore, in these cards, we reported coordinates according to different GenBank accession numbers. For example, the HIV-1 genome has been mainly referred to K03455 and AF033819 NCBI IDs but their coordinates are shifted of 455 nt, so we reported both notations.

4 IMPLEMENTATION

For storage and query of the data, we constructed a MySQL relational database maintained in our server. In the main table, for each sequence, we annotated the species it belongs to, the name of harboured functional elements, their synonyms, details about the structure of the investigated sequence or the possible overlapped genes, cell lines used to assess the functional effect, genomic coordinates, a score describing the functional effect and the related literature.

For each record, we created an image showing the genomic context the sequence derives from. The second table stores the *trans*-acting factors and the involved export pathways.

We also created a user-friendly web interface that allows users to browse RNA sequences and align a sequence against ExportAid collected sequences. The web interface uses Perl DBI and CGI, Javascript, and Ajax/XML languages to connect with the database, facilitate data visualization and retrieval. Moreover, the wide use of cross links to external databases allows an easy switch to source database, such as NCBI taxonomy (organism field), NCBI EntrezGene (sequence structure field), NCBI nucleotide (coordinates field) and PubMed (references field).

5 DISCUSSION

To our knowledge, ExportAid is the only resource that collects information about the RNA elements regulating nuclear export. All of them have been experimentally assessed in mammalian contexts and belong to both mammalian and viral genomes while others are artificial. We also stored sequences that have no effect on export. This comprehensive collection of experimental data can be of great interest, for example, to build binding preference models of RNA export proteins to be used in prediction algorithms. Using neutral sequences, as examples to train machine learning algorithms, it is possible to obtain prediction tools with a low rate of false-negative and -positive results. This work highlights which functional elements have been poorly investigated and should be further studied to build binding preference models.

This database could help to understand the rules of the RNA export language and to build an algorithm to predict the fate of gene transcripts, non-coding RNAs and viral RNAs, both normal or mutated. In particular, prediction of molecular effects of polymorphisms constitutes the core of the prioritization process that is carried out before association studies (Blasi *et al*., 2013; Galeazzi *et al*., 2014; Piva *et al*., 2011). Moreover, ExportAid could be used in virology to study the effect of emerging viral mutants, or to design lentiviral gene therapy vectors with more efficient RNA export elements (Oh *et al*., 2007).

Our data could also help to explain the export or the retention of transcripts, such as intronless genes, some lncRNAs and viral genomes, that do not undergo splicing and so do not have the canonical adapter proteins. Similarly, it may be useful to understand why some transcripts, despite undergoing splicing, are retained in the nucleus. For example, the steroid receptor RNA activator (SRA) lncRNA is processed by splicing machinery, but its mature sequence is retained in the nucleus. This retention is not because of adenosine-to-inosine editing, as SRA stem structures are shorter than 25/30 bp, the minimum length required to be edited and thus retained in the nucleus (DeCerbo and Carmichael, 2005; Novikova *et al*., 2012, 2013). Analogously, the retention of the Gomafu lncRNA has not been explained; in fact, it was hypothesized that it was owing to binding with SF1 but subsequently this involvement was not confirmed (Sone *et al*., 2007; Tsuiji *et al*., 2011). Also the X-inactive-specific

transcript (XIST) lncRNA is retained in the nucleus after intron removal but this seems due to inefficient binding with the EJC protein Magoh and the export factor TAP/NXF1 (Cohen and Panning, 2007). However, the lncRNAs can have an active role in regulating the RNA export. For example, the reduction of NEAT1 impairs the editing-associated nuclear retention pathway (Chen and Carmichael, 2009) and even its knockdown increases the nucleus-to-cytoplasm export of Rev-dependent HIV-1 transcripts (Zhang *et al.*, 2013).

Our data can also be useful to understand the mRNA export dysregulation in cancer. This is because of aberrant expression of export factors like, for example, the eukaryotic translation initiation factor eIF4E (Carroll and Borden, 2013) that increases the export of specific transcripts involved in proliferation, survival and oncogenesis (Culjkovic-Kraljacic and Borden, 2013). Besides the critical role in cancer genesis and maintenance, RNA export can be modulated in response to cellular signalling or stress, such as growth factors, cytokines, chemokines and oncogenic tyrosine kinases that are able to induce THOC5, a member of the TREX export complex (Tran *et al.*, 2014). In addition, it was recently reported that, GANP protein, subunit of the TREX complex, promotes the nuclear export of a subset of mRNAs involved in RNA synthesis and processing, thus facilitating rapid changes in expression of many genes (Wickramasinghe *et al.*, 2014). Unfortunately, it is not yet known which export motifs, lying in these mRNAs, bind GANP.

In this work, we realized that different splicing proteins (9G8, SRp20, SF2/ASF, hnRNP A2/B1, hnRNP C, hnRNP L and HuR) are also involved in RNA export suggesting a crosstalk between these two processes. According to SpliceAid 2 (Piva *et al.*, 2012) resource, these splicing proteins recognize degenerate RNA motifs that are spread in mRNA and not only at exonic junctions, prompting export language to be overlapped both to the splicing language and the genetic code.

Interestingly, whereas splicing language is fragile, as most mutations affect the correct intron removal giving rise to aberrant splicing patterns, the export language seems to be robust in genes that undergo splicing. In fact, the export of only two mRNAs (collagen COL1A1 and dystrophia myotonica-protein kinase, DMPK) is impaired by mutations (Hurt and Silver, 2008) unlike intronless genes (Grzybowska, 2012).

Finally, we plan to continuously update the database, also with the contribution of users who can signal new elements.

ACKNOWLEDGEMENT

The authors acknowledge a kind thanks to Mrs Monica Glebocki for language editing.

Conflict of interest: none declared.

REFERENCES

Anko,M.L. (2014) Regulation of gene expression programmes by serine-arginine rich splicing factors. *Semin. Cell Dev. Biol.*, **32**, 11–21.

Barbieri,M. (2003) *The Organic Codes: An Introduction to Semantic Biology*. Cambridge University Press, Cambridge, UK.

Blasi,G. *et al.* (2013) Association of GSK-3beta genetic variation with GSK-3beta expression, prefrontal cortical thickness, prefrontal physiology, and schizophrenia. *Am. J. Psychiatry*, **170**, 868–876.

Carmody,S.R. and Wente,S.R. (2009) mRNA nuclear export at a glance. *J. Cell. Sci.*, **122**, 1933–1937.

Carroll,M. and Borden,K.L. (2013) The oncogene eIF4E: using biochemical insights to target cancer. *J. Interferon. Cytokine Res.*, **33**, 227–238.

Chen,L.L. and Carmichael,G.G. (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell*, **35**, 467–478.

Cohen,H.R. and Panning,B. (2007) XIST RNA exhibits nuclear retention and exhibits reduced association with the export factor TAP/NXF1. *Chromosoma*, **116**, 373–383.

Colosimo,M.E. *et al.* (2005) Data preparation and interannotator agreement: BioCreAtIvE task 1B. *BMC Bioinformatics*, **6** (Suppl. 1), S12.

Culjkovic-Kraljacic,B. and Borden,K.L. (2013) Aiding and abetting cancer: mRNA export and the nuclear pore. *Trends Cell Biol.*, **23**, 328–335.

Culjkovic,B. *et al.* (2006) eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J. Cell. Biol.*, **175**, 415–426.

Cullen,B.R. (2003) Nuclear mRNA export: insights from virology. *Trends Biochem. Sci.*, **28**, 419–424.

DeCerbo,J. and Carmichael,G.G. (2005) Retention and repression: fates of hyperedited RNAs in the nucleus. *Curr. Opin. Cell Biol.*, **17**, 302–308.

Donello,J.E. *et al.* (1998) Woodchuck hepatitis virus contains a tripartite posttranscriptional regulatory element. *J. Virol.*, **72**, 5085–5092.

Erkmann,J.A. and Kutay,U. (2004) Nuclear export of mRNA: from the site of transcription to the cytoplasm. *Exp. Cell. Res.*, **296**, 12–20.

Galeazzi,R. *et al.* (2014) Insights into the influence of 5-HT2c aminoacidic variants with the inhibitory action of serotonin inverse agonists and antagonists. *J. Mol. Model*, **20**, 2120.

Giulietti,M. *et al.* (2013) SpliceAid-F: a database of human splicing factors and their RNA-binding sites. *Nucleic Acids Res.*, **41**, D125–D131.

Grzybowska,E.A. (2012) Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochem. Biophys. Res. Commun.*, **424**, 1–6.

Higashimoto,T. *et al.* (2007) The woodchuck hepatitis virus post-transcriptional regulatory element reduces readthrough transcription from retroviral vectors. *Gene Ther.*, **14**, 1298–1304.

Hurt,J.A. and Silver,P.A. (2008) mRNA nuclear export and human disease. *Dis. Model Mech.*, **1**, 103–108.

Kohler,A. and Hurt,E. (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.*, **8**, 761–773.

Le Hir,H. *et al.* (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.

Leblanc,J. *et al.* (2013) Posttranscriptional regulation of retroviral gene expression: primary RNA transcripts play three roles as pre-mRNA, mRNA, and genomic RNA. *Wiley Interdiscip. Rev. RNA*, **4**, 567–580.

Lei,H. *et al.* (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.

Lei,H. *et al.* (2013) Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export. *Nucleic Acids Res.*, **41**, 2517–2525.

Novikova,I.V. *et al.* (2012) Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res.*, **40**, 5034–5051.

Novikova,I.V. *et al.* (2013) Rise of the RNA Machines: exploring the Structure of Long Non-Coding RNAs. *J. Mol. Biol.*, **425**, 3731–3746.

Oh,T. *et al.* (2007) Lentiviral vector design using alternative RNA export elements. *Retrovirology*, **4**, 38.

Palazzo,A.F. *et al.* (2007) The signal sequence coding region promotes nuclear export of mRNA. *PLoS Biol.*, **5**, e322.

Piva,F. *et al.* (2011) Bioinformatic analyses to select phenotype affecting polymorphisms in HTR2C gene. *Hum. Psychopharmacol.*, **26**, 365–372.

Piva,F. *et al.* (2012) SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. *Hum. Mutat.*, **33**, 81–85.

Prechtel,A.T. *et al.* (2006) Expression of CD83 is regulated by HuR via a novel cis-active coding region RNA element. *J. Biol. Chem.*, **281**, 10912–10925.

Rodrigues,J.P. *et al.* (2001) REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. *Proc. Natl Acad. Sci. USA*, **98**, 1030–1035.

Sone,M. *et al.* (2007) The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *J. Cell Sci.*, **120**, 2498–2506.

Stewart,M. (2010) Nuclear export of mRNA. *Trends Biochem. Sci.*, **35**, 609–617.

Tran,D.D. *et al.* (2014) THOC5, a member of the mRNA export complex: a novel link between mRNA export machinery and signal transduction pathways in cell proliferation and differentiation. *Cell Commun. Signal.*, **12**, 3.

Tsuiji,H. *et al.* (2011) Competition between a noncoding exon and introns: gomafu contains tandem UACUAAC repeats and associates with splicing factor-1. *Genes Cells*, **16**, 479–490.

Wickramasinghe,V.O. *et al.* (2014) Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic Acids Res.*, **42**, 5059–5071.

Zhang,Q. *et al.* (2013) NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. *MBio*, **4**, e00596–e00512.

Zhou,Z. *et al.* (2000) The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature*, **407**, 401–405.

Zufferey,R. *et al.* (1999) Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J. Virol.*, **73**, 2886–2892.