

# Implications of polyadenylation in health and disease

Ana Curinha<sup>1,2,†</sup>, Sandra Oliveira Braz<sup>1,2,†</sup>, Isabel Pereira-Castro<sup>1</sup>, Andrea Cruz<sup>1</sup>, and Alexandra Moreira<sup>1,\*</sup>

<sup>1</sup>Gene Regulation Group; IBMC-Instituto de Biologia Molecular e Celular; Universidade do Porto; Porto, Portugal; <sup>2</sup>Departamento de Biologia; Universidade de Aveiro; Aveiro, Portugal

<sup>†</sup>These authors contributed equally to this work.

**Keywords:** gene expression, polyadenylation, alternative polyadenylation, 3' untranslated region, disease, cell state

**Abbreviations:** APA, Alternative polyadenylation; pA tail, polyA tail; PAP, polyA polymerase; PAS, polyA site; pA signal, polyA signal; UTR, untranslated region; miRNA, microRNA; RBP, RNA-binding protein; RNA Pol II, RNA polymerase II; TLI, tandem UTR length index; IMP-1, Insulin-like growth factor 2 mRNA binding protein 1; 3'READS, 3' Region Extraction and Deep Sequencing; DMKN, dermokine; PDXK, pyridoxal kinase; PPIE, peptidylpropylisomerase E; CFIm25, Cleavage Factor Im 25 kDa;  $\mu$ , IgM heavy-chain mRNA; CSTF2, cleavage stimulatory factor-64kDa; DSE, downstream sequence element; siRNAs, small interfering RNAs; TCR, T cell receptor; LPS, lipopolysaccharide; TNF- $\alpha$ , tumor necrosis factor- $\alpha$ ; AREs, Au-rich elements; IPEX, immune dysfunction, polyendocrinopathy, enteropathy, X-linked; FOXP3, forkhead box P3; WAS, Wiskott-Aldrich syndrome; WASP, Wiskott-Aldrich syndrome protein; USE, upstream sequence element; CPSF, Cleavage and Polyadenylation Specificity Factor; ESC, embryonic stem cells; SLE, systemic lupus erythematosus; SNP, single nucleotide polymorphism; COX-2, cyclooxygenase 2; AD, Alzheimer disease; aSyn,  $\alpha$ -Synuclein; PD, Parkinson disease; aSynL, longest aSyn isoform; OPMD, oculopharyngeal muscular dystrophy; PABPN1, poly(A) binding protein; nuclear 1;  $\alpha$ -GalA,  $\alpha$ -galactosidase A; FMR1, Fragil X mental retardation 1; FXS, Fragile X syndrome; FXPOI, fragile X-associated immature ovarian insufficiency; FXTAS, fragile X-associated tremor/ataxia syndrome; snRNPs, spliceosomal small nuclear ribonucleoproteins; BPV, bovine papilloma virus; SMN, Survival Motor Neuron; SMA, Spinal Muscular Atrophy; StAR, steroidogenic acute regulatory; CAH, congenital adrenal hyperplasia; HGRG-14, high-glucose-regulated gene; TCF7L2, transcription factor 7-like 2; TCF/LEF, T cell factor/lymphoid enhancer factor.

Polyadenylation is the RNA processing step that completes the maturation of nearly all eukaryotic mRNAs. It is a two-step nuclear process that involves an endonucleolytic cleavage of the pre-mRNA at the 3'-end and the polymerization of a polyadenosine (polyA) tail, which is fundamental for mRNA stability, nuclear export and efficient translation during development. The core molecular machinery responsible for the definition of a polyA site includes several recognition, cleavage and polyadenylation factors that identify and act on a given polyA signal present in a pre-mRNA, usually an AAUAAA hexamer or similar sequence. This mechanism is tightly regulated by other *cis*-acting elements and *trans*-acting factors, and its misregulation can cause inefficient gene expression and may ultimately lead to disease. The majority of genes generate multiple mRNAs as a result of alternative polyadenylation in the 3'-untranslated region. The variable lengths of the 3' untranslated regions created by alternative polyadenylation are a recognizable target for differential regulation and clearly affect the fate of the transcript, ultimately modulating the expression of the gene. Over the past few years, several studies have highlighted the importance of polyadenylation and alternative polyadenylation in gene expression and their impact in a variety of physiological conditions, as well as in several illnesses. Abnormalities in the 3'-end processing

mechanisms thus represent a common feature among many oncological, immunological, neurological and hematological disorders, but slight imbalances can lead to the natural establishment of a specific cellular state. This review addresses the key steps of polyadenylation and alternative polyadenylation in different cellular conditions and diseases focusing on the molecular effectors that ensure a faultless pre-mRNA 3' end formation.

## Introduction

Gene expression is a biological process essential to all organisms that is tightly regulated at different steps. The regulation of gene expression is accomplished through modulation of the chromatin structure by the interaction of regulatory proteins with specific DNA motifs of the target genes and also by RNA processing. At the pre-mRNA processing level that occurs in the nucleus, gene expression is mainly modulated by alternative splicing and alternative polyadenylation (APA), which have emerged as key regulatory mechanisms that allow the production of many different transcripts from one single transcription unit.<sup>1-3</sup>

The 3'-ends of nearly all eukaryotic mRNAs are formed by cleavage and polyadenylation, which are necessary for nuclear export and stability of the mature transcripts. This two-step nuclear process involves an endonucleolytic cleavage of pre-mRNA, followed by the polymerization of a polyA (pA) tail to

\*Correspondence to: Alexandra Moreira; Email:alexandra.moreira@ibmc.up.pt  
Submitted: 06/06/2014; Revised: 08/30/2014; Accepted: 09/04/2014  
<http://dx.doi.org/10.4161/nucl.36360>

**Table 1.** Summary of polyadenylation (PA) and alternative polyadenylation (APA) related diseases

Conditions	Condition and/or Disease	Gene	PA/APA	Mechanism	Methodology	Ref.
Hemathological	$\beta$ -thalassemia	$\beta$ -globin	APA	Mutations on pA cause elongated 3'UTR and transcription termination defects	RNase protection, northern blotting	17, 18
	$\alpha$ -thalassemia	$\alpha$ -globin	APA		Southern blotting, SSSA analysis	19, 20
	Thrombophilia	<i>Prothrombin</i>	PA	Increase of cleavage site recognition due to a G20210A variation on PAS	Reporter assay, RT-qPCR, Western blotting	24, 26
Proliferation and Oncological	Proliferative conditions	—	APA	Shortening of 3'UTR due to proximal PAS usage	PLATA, 3' deep sequencing technique	16, 29
	Cancer	E.g.: <i>IMP-1</i>	APA	Shortening of 3'UTR due to proximal PAS usage	High-throughput sequencing, Reporter assays, western blotting	15, 33
	Breast cancer (MB231 cell line)	—	APA	Lengthening of 3'UTR due to proximal PAS usage	3'-deep sequencing, 3'RACE, RT-qPCR	31
	Colorectal cancer	<i>DMKN, PDXK</i> and <i>PPIE</i>	APA	Shortening of 3'UTR during cancer progression	3'-sequencing, 3'PCR, RT-qPCR	35
	Glioblastoma	<i>CCND1</i> and <i>MECP2</i>	APA	Shortening of 3'UTR due to CFIm25 depletion	RNA-Seq, western blotting, RT-qPCR	36
Infection and Immunological	B-cell differentiation	<i>IgM heavy-chain</i> ( $\mu$ )	APA	CSTF2 promotes weaker proximal PAS usage	Western blotting, in vitro cleavage reactions	41
				U1A regulates pA tail addition to $\mu$ secretory PAS	EMSAs, pA addition assays, pA tail measurements	42
				U1A inhibits CSTF2 binding to $\mu$ secretory PAS	EMSAs, UV-crosslinking assays, cleavage assays	37, 43
	T-cell activation	<i>NF-ATc</i>	APA	CTD tail phosphorylation, ELL2 and histone methylation promote the usage of proximal PAS	ChIP, RT-qPCR	14, 37
				Shortening of 3'UTR upon T-cells activation due to increased levels of CSTF2	Reporter assay, northern blotting, RNase protection assay, RNA EMSAs, western blotting	45
				Shortening of 3'UTR due to increased levels of CSTF2	Western blotting, RNase protection assay, RT-qPCR	46
	LPS stimulation of macrophages in infection	—	APA	Shortening of 3'UTR due to increased levels of CSTF2	Western blotting, RNase protection assay, RT-qPCR	46
	Inflammation and host defense	<i>TNF-<math>\alpha</math></i>	PA	Translation rate decrease due to pA tail absence in LPS-stimulated macrophages	Northern blotting, RT-qPCR, Polysome analysis, DNA microarray	48, 51, 52
	IPEX syndrome	<i>FOXP3</i>	PA	Mutations in pA signal (AATAAA - AATGAA)	3' RACE, RT-qPCR	54, 56
	Wiskott-Aldrich syndrome	<i>WASP</i>	APA	Mutations on 3'UTR that introduce new pA signals	3' RACE, RT-qPCR, western blotting	60
Osteoarthritis and Scleroderma	<i>Collagen</i> genes	PA	USE-dependent polyadenylation	In vitro transcription, polyadenylation and cleavage assays	63	
Systemic Lupus Erythematosus	<i>GIMAP5</i>	APA	Mutation on proximal pA signal (AATAAA - AATAGA)	Western blotting, northern blotting, RT-qPCR	72	
Neurological	Alzheimer Disease	<i>COX-2</i>	APA	Increased distal PAS usage	Northern blotting, RT-qPCR	84
			APA	Increase in dopamine levels leads to distal PAS usage	PolyA RNA-Seq, Reporter Assays, Western and northern blotting	88
	Oculopharyngeal muscular dystrophy	<i>PABPN1</i>	APA	PABPN1 mutation leads to 3'UTR shortening	Reporter assay, Co-immunoprecipitation, R-EMSA, Cleavage assay	94, 96

(continued on next page)

**Table 1.** Summary of polyadenylation (PA) and alternative polyadenylation (APA) related diseases (Continued)

Conditions	Condition and/or Disease	Gene	PA/APA	Mechanism	Methodology	Ref.
	Fabry disease	<i>α-GalA</i>	APA	mRNA instability due to mutations in pA signal	3' RACE, northern blotting	104
	Fragile X syndrome	<i>FMR1</i>	APA	Repeats on 5' UTR affect choice of pA signals	RLM-RACE, western blotting, polyadenylation test, polysome/mRNPs gradient	109
	Spinal Muscular Atrophy	<i>SMN</i>	PA	U1A inhibits cleavage and polyadenylation	EMSA, cleavage and polyadenylation assays, Reporter assay	111
Endocrine	Steroidogenesis	<i>StAR</i>	APA	Br-cAMP promotes distal PAS usage	Northern blotting, western blotting, Reporter assay	120
	Diabetic Nephropathy	<i>HGRG-14</i>	APA	Increased levels of glucose lead to distal PAS usage	mRNA differential display, northern blotting, transcription stability	124
	Type II diabetes	<i>TCF7L2</i>	APA	Increased usage of intronic PAS	Western blotting, reporter assay	130

the mRNA molecule by polyA polymerase (PAP). The length of the pA tail is crucial for the transport of the mature mRNAs to the cytoplasm, their translation efficiency in certain developmental stages, and the quality control and degradation of mRNA. Although it was initially established that most mRNAs contained a pA tail of ~150–250 adenosines<sup>1–5</sup> recent studies have now challenged this view. Two different deep sequencing methods were recently developed to show that the mean pA tail length is generally much shorter: ~30 nucleotides in yeast and ~50–100 nucleotides in mammalian and *Drosophila* cell lines.<sup>6,7</sup> Surprisingly, a strong correlation between the length of the pA tail and translation efficiency could only be found in early embryos. The pA tail is thus a dynamic region of the mRNA<sup>8</sup> that is controlled differently depending on a specific developmental stage. Interestingly, it has been shown that an increase in PAP activity is associated with poor prognosis in certain cancers<sup>9</sup> and also that usage of PAP inhibitors affects some genes involved in inflammatory conditions.<sup>10</sup> Therefore, a tight control of the pA tail length may be a determinant factor in the development of some diseases.

Cleavage and polyadenylation occur co-transcriptionally, and in addition to transcription per se, the main determinant factors responsible for the definition of a specific polyA site (PAS) in the pre-mRNA include *cis*-acting RNA elements and several core and auxiliary *trans*-acting factors that participate and enhance the efficiency of the process. The most important *cis*-element is a 6 nucleotide hexamer (AAUAAA), the polyA signal (pA signal), located ~10–35 nucleotides upstream of the PAS. Although the AAUAAA sequence is the canonical pA signal, the pA signal can adopt more than ten weaker variants of that sequence.<sup>1–5,11</sup>

APA is a widespread mechanism in eukaryotic cells that controls gene expression by producing mRNA isoforms with alternative 3' ends. Usage of one PAS over another is often attributed to the relative strength of the *cis* and *trans*-acting elements mentioned above but a coordinated integration with transcription has to take place.<sup>12–14</sup>

The different types of APA described previously<sup>1</sup> differ in the localization of the pA signal. When the cleavage and/or polyadenylation machinery uses a pA signal localized in introns or exons, it produces mRNAs with different coding sequences resulting in different proteins. However, when it recognizes pA signals located in the 3' untranslated region (UTRs), it creates mRNAs of different lengths. As 3' UTRs often have binding sites for microRNAs (miRNAs) and RNA-binding proteins (RBPs), naturally longer 3' UTRs contain more of these regulatory sequences than shorter 3' UTRs. Consequently, the differential utilization of one or another pA signal in the 3'UTR implies that those *cis*-elements will be differentially present in the mRNA, affecting the fate of the transcript and ultimately modulating the expression of the gene.<sup>1–4,11,15,16</sup>

Polyadenylation and APA have garnered much interest in recent years as several studies have revealed a crucial role for this event in the control of different cellular processes such as proliferation, differentiation and development. Importantly, it has been shown that abnormalities in the 3'-end processing mechanism are a common feature among many oncological, immunological, neurological, and hematological diseases, as well as in cellular and molecular conditions important for cell homeostasis (Table 1).<sup>1–4,11,15,16</sup> This review discusses some examples where polyadenylation and APA play fundamental roles in the establishment of diseases and also in different cellular conditions.

## Diseases and Cellular States Associated to Altered Polyadenylation

### Hematological disorders

Alterations in one of the most important sequence elements present in 3'UTR, the pA signal, have been linked to numerous hematological diseases. The majority of these studies were initially conducted in thalassemias, a group of inherited autosomal

recessive hematological disorders caused by defects in the synthesis of one or more of the hemoglobin chains. These studies clearly showed that mutations on a pA signal severely affect the expression of globin genes and are responsible for the disease. These mutations lead to failure in a globin pA signal usage and as RNA polymerase II (RNA Pol II) transcription termination mostly depends on the pA signal, elongated mRNA isoforms were produced. This causes changes in protein expression and contributes to thalassemia phenotypes. Examples of early works on these mutations in globin pA signals and their implications in  $\alpha$  and  $\beta$ -thalassemia are described below.

In 1985 a point mutation in the canonical pA signal (AATAAA to AACAAA) of a human  $\beta$ -globin gene isolated from a patient with  $\beta$ -thalassemia was identified. This mutation leads to the formation of an elongated  $\beta$ -globin mRNA isoform using a canonical pA signal 900 nucleotide downstream of the mutation site.<sup>17</sup> Later, two other mutations in the  $\beta$ -globin pA signal were identified, a deletion (AATAAA to A—) and a point mutation (AATAAA to AATAAG). The point mutation leads to the production of four new elongated mRNA isoforms, which causes a decrease in  $\beta$ -globin mRNA expression to 12–34%.<sup>18</sup>

In the case of  $\alpha$ -thalassemias, a single substitution in the pA signal of  $\alpha_2$ -globin (AATAAA to AATAAG) was identified in 1983 and it was discovered that the expression of the downstream  $\alpha_1$ -globin gene was inactivated.<sup>19</sup> Later on, it was identified a mutation on the pA signal of the  $\alpha_2$ -globin gene characterized by the deletion of two base pairs (AATAAA to AATA—) that causes a decrease in the expression of  $\alpha_2$ -globin gene. As this mutation affects RNA Pol II transcription termination, expression of  $\alpha_1$ -globin gene was also affected by transcription interference.<sup>20</sup> These studies highlight the importance of pA signals, since their disruption not only affects the encoded gene but also may interfere with the transcription of other genes, as in the case of  $\alpha_2$ - and  $\alpha_1$ -globin.

Several studies have disclosed the role of prothrombin pre-mRNA processing in thrombophilia and the mechanisms involved in that process. It was found that a G20210A change at prothrombin PAS is responsible for increased levels of prothrombin in the plasma<sup>21–26</sup> and is a risk factor for venous thrombosis.<sup>27</sup> Although the G20210A variation does not affect the site where cleavage and polyadenylation occur or the length of the pA tail itself, it leads to an increase in the efficiency of pre-mRNA cleavage. The optimal site of pre-mRNA cleavage and polyadenylation is usually a CA but in pro-thrombin the PAS is CG, which represents an inefficient cleavage site. The variation G20210A changes the G to an A, resulting in an increase in cleavage and polyadenylation. Indeed, using constructs containing 20210A or 20210G sequences in tandem, Ceelie et al. 2003 demonstrated that transcription tends to terminate more efficiently at the 20210A site.<sup>26</sup> This causes an increase in the 3' end processing of prothrombin pre-mRNA and consequently in the production of higher levels of protein.<sup>24</sup> Therefore, the G20210A change in the prothrombin gene represents a paradigm of polyadenylation gain-of-function leading to a severe risk of thrombosis development.

### Proliferation and oncological conditions

Following recent advances in genome wide high-throughput methodologies, several studies have demonstrated the tendency of proliferative and tumor cells to produce mRNAs with shorter 3' UTRs, eluding miRNAs and RBPs regulation. Interestingly, it has also been shown that during development and differentiation the opposite occurs: there is a preferential choice of distal PAS,<sup>28</sup> resulting in mRNA isoforms with longer 3'UTRs.

In T lymphocytes, Sandberg et al. 2008 developed a tandem UTR length index (TLI) that assessed the expression of extended 3'UTR relative to total gene expression levels to show that in CD3/CD28 stimulated mouse CD4<sup>+</sup>T cells, human B-cells and monocytes, the TLI is decreased after 48h. This study therefore showed a general decrease in the relative expression of distal pA signal in tandem UTRs upon T cell activation. Additionally, it established a negative correlation between proliferative index and TLI as TLI values were lower in cell lines than in normal tissues from which the respective cell line derived. These results further suggest that under proliferative conditions, 3'UTRs are generally shorter.<sup>16</sup> As B and T lymphocyte activation is a key immune response to several stimuli, this may represent an efficient mechanism to escape regulation and respond quicker when lymphocyte activation is necessary.

Remarkably, a similar observation was reported in cancer cells. Although cancer and non-transformed cell lines have similar proliferation rates, the former tend to produce increased levels of mRNA isoforms with shorter 3'UTR. This reveals that the correlation between shorter 3'UTRs and cell transformation is higher than the correlation of shorter 3' UTRs and cell proliferation. It was also shown that globally the shortest mRNA isoforms have higher stability and produced more protein, which may be a relevant activation mechanism employed by some oncogenes in cancer cells. Indeed, the shortest mRNA isoform of the proto-oncogene insulin-like growth factor 2 mRNA binding protein 1 (IMP-1) was shown to promote higher oncogenic transformation levels than its longest mRNA isoform.<sup>15</sup>

The increase in mRNAs with shorter 3'UTRs observed in proliferative conditions<sup>16</sup> was also confirmed in primary fibroblast BJ cells and non-transformed MCF10A cells.<sup>29</sup> It was reported that in proliferative conditions, elevated levels of E2F proteins lead to an increase in the expression of 3' end processing genes<sup>30</sup> through binding to their promoter. It was therefore suggested that this caused the increase in cleavage and polyadenylation at the proximal PAS observed in proliferative cells.<sup>29</sup>

Notwithstanding these results, in two different breast cancer cell lines (MCF7 and MB231) an opposite pattern of 3'UTR length was also described. According to what had been previously described in cancer and proliferative conditions, the MCF7 cell line demonstrated a high production of shorter 3'UTR mRNA isoforms; however, the MB231 cell line showed an opposite switch to longer 3'UTRs.<sup>31</sup> This demonstrates that in cancer APA is cell type-dependent, as has recently been shown in healthy cellular conditions.<sup>32</sup> Consequently, the cellular states of proliferation and transformation are not the only determinants involved in this regulatory mechanism. Additionally, regulation by

miRNAs has an important role in decreasing the expression of elongated 3'UTRs, as previously described by Sandberg et al. 2008 for activated T cells.<sup>16</sup> In MCF7 and MCF10A cancer cell lines in particular, it was found that miRNA-25/32/92/92ab/363/367 have a role in decreasing the ratio between the long and short mRNA isoforms of its target genes.<sup>33</sup>

As opposed to cell proliferation, Tian and coll. have shown that during embryonic development and cell differentiation there is a global increase in distal PAS usage. The authors developed the method 3' Region Extraction And Deep Sequencing (3'READS), which diminishes the issue of internal priming and oligo(A) tail in PAS mapping. They analyzed C2C12 and 3T3-L1 cells, which were induced to differentiate and represent a model for myogenesis and adipogenesis respectively, and also mouse embryos as a model for embryogenesis. A dramatic upregulation of mRNAs produced by distal PAS was observed both in mouse embryos and during cell differentiation. Interestingly, it was also observed that upregulated mRNA isoforms present generally stronger pA signals, indicating that the strength of a pA signal is an important determinant in PAS selection during development and differentiation.<sup>34</sup>

During colorectal cancer development, alterations in APA patterns were observed during the different stages of cancer progression: normal mucosa, adenoma, and carcinoma. APA alterations of three genes were characterized: dermokine (*DMKN*), pyridoxal kinase (*PDXK*), and peptidylpropyl isomerase E (*PPIE*). A switch to proximal PAS usage was found for *DMKN* and *PPIE* during progression from normal mucosa to adenoma, and from normal mucosa to carcinoma. For the *PDXK* and *PPIE* genes it was also shown that the APA pattern was altered during progression from adenoma to carcinoma. The mechanism behind this observation probably involves an increase in the expression of some components of the polyadenylation machinery detected during colorectal cancer progression (from normal mucosa to carcinoma). This would induce usage of the proximal PAS, which is generally more inefficient than the distal one, in a similar mechanism observed for other genes.<sup>35</sup>

It has been steadily demonstrated that alterations in the levels of cleavage and polyadenylation factors in different cellular conditions affect the APA pattern observed. A recent study on glioblastoma cell lines demonstrated the relevance of the cleavage factor CFIm25 in this process in tumorigenesis. After CFIm25 knockdown a global switch to the usage of the proximal PAS was observed, suggesting a function of this *trans*-acting factor as a repressor of proximal PAS selection. In particular, CFIm25 knockdown causes an increase in the levels of several proteins encoded by tumor growth-related genes, including cyclin D1, concomitant with a switch to the proximal PAS. Taking advantage of the different levels of CFIm25 in two different glioblastoma cell lines (U251 that express low levels of CFIm25 and LN229 that express high levels of CFIm25), it was shown that overexpression of CFIm25 in U251 leads to a decrease in cell growth and cellular invasion, and downregulation of CFIm25 in LN229 lead to an increase in tumorigenic properties. These results suggest a role for CFIm25 as a potential tumor suppressor through regulation of APA events.<sup>36</sup>

Taken together, these studies indicate clearly that APA is not regulated by a single factor, but it is regulated in a cell and/or tissue-specific manner. It is conceivable that the shift to the proximal PAS is a mechanism to evade regulation that could hamper cell cycle progression during cell activation, proliferation, and in some cancers. Nevertheless, all these studies show that alterations in cleavage and/or polyadenylation factors have a dramatic impact in the cellular state due to major changes in APA and consequently in gene expression.

### Infection and immunological conditions

Immune responses are essential to control infectious and immunological diseases by inducing inflammatory reactions. This requires a coordinated and timely regulated cellular response involving multiple genes. Any alteration in this regulation has an enormous effect in the magnitude and duration of the inflammatory reaction and can lead to pathological conditions.

As reviewed previously<sup>37,38</sup> a variety of studies have assessed how immunoglobulin heavy chain genes are regulated at the RNA processing level during B-cell differentiation. The IgM heavy-chain mRNA ( $\mu$ ) undergoes alternative processing, involving competition of splicing and polyadenylation events, to produce two distinct transcripts that encode a membrane receptor or a secreted antibody during B-cell differentiation. In pre-B and B cells, the pA signal of the secretory mRNA isoform is spliced out resulting in the production of the  $\mu$  membrane receptor, while in differentiated plasma cells, the proximal secretory PAS is used, which results in the production of the mRNA encoding the  $\mu$  secreted form.<sup>39,40</sup> The mechanism involved in the usage of one pA signal over another involves a variety of mechanisms implicating RBPs and *cis*-sequence elements, which play important roles on defining which PAS is used and which  $\mu$  mRNA isoform is produced during B-cell development. It has been shown that variations in the expression of Cleavage Stimulatory Factor-64 kDa (CSTF2), which binds to a GU/U-rich *cis*-element downstream to the PAS (DSE), affect  $\mu$  pre-mRNA processing. B-cells express decreased levels of CSTF2 and the  $\mu$  membrane pA signal is more used, while in plasma cells, that express increased levels of CSTF2, the weaker  $\mu$  secretory pA signal is recognized more efficiently and thus used.<sup>41</sup> U1A was also shown to affect  $\mu$  pre-mRNA APA. Three U1A sequence motifs localized upstream of the  $\mu$  secretory pA signal regulate pA tail addition to the  $\mu$  secretory mRNA and control its expression during B-cell differentiation.<sup>42</sup> Two other downstream U1A-binding motifs inhibit the binding of CSTF2 to its natural binding site in the pre-mRNA and thus are also implicated in the process, impeding cleavage at the  $\mu$  secretory PAS.<sup>43</sup>

More recently it has been shown that the choice between  $\mu$  secretory and membrane isoforms PAS also depends on transcription elongation, as the ELL2 elongation factor promotes the production of the secretory mRNA isoform.<sup>14,44</sup> It was also shown that ELL2 knockdown causes a decrease in histone 3 methylation (H3K4 and H3K79) and as B cells and plasma cells present different levels of H3K4 and H3K79 methylation, it is possible that these chromatin marks also play a role in *IgH* RNA processing. These results suggest that an increased transcription in plasma

cells facilitates the recognition of the proximal PAS and the splicing event characteristic of secretory isoform production.<sup>44</sup> Additionally, they indicate a role for RNA Pol II and transcription-related factors in *IgH* RNA processing.

The mRNA isoforms of the transcription factor NF-ATc are differentially regulated during T cell activation in a similar manner. Three different NF-ATc transcripts are produced due to splicing and APA events, differing both in the coding region and also in the 3' UTR length. In naïve T cells, two longer isoforms are produced (NF-ATc/B and NF-ATc/C), whereas in effector T cells a weak proximal pA signal is used, resulting in a shorter mRNA isoform (NF-ATc/A). It was also shown that CSTF2 expression is increased upon T cell receptor (TCR) stimulation. As in naïve T cells CSTF2 levels are lower, the proximal pA signal is not efficiently recognized which allows the splice event to occur and the distal PAS to be used. Upon TCR activation, CSTF2 expression levels increase, leading to efficient use of the proximal pA signal. This is another example where splicing and polyadenylation competition play a role in pA site choice in the immune system cells.<sup>45</sup> Interestingly, macrophages stimulated with lipopolysaccharide (LPS) show an increase in CSTF2 expression which also leads to an increase in proximal PAS selection for several mRNAs.<sup>46</sup> Therefore, alterations in the levels of this important cleavage and/or polyadenylation factor may affect pre-mRNA processing in several cellular states.

Tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) is a key player in inflammation and host defense.<sup>47</sup> Studies on TNF- $\alpha$  mRNA regulation<sup>48</sup> have demonstrated that while in unstimulated macrophages TNF- $\alpha$  mRNA is not translated, upon activation there is an increase in its translation rate. This is due to a variety of *cis*-acting elements present on TNF- $\alpha$  3'UTR<sup>48</sup> that surround the UUAUUUUAU sequence, known as AU-rich elements (AREs), which repress translation by recruiting deadenylases and downstream degradation machineries.<sup>49,50</sup> Interestingly, northern blot analysis shows that translational silent TNF- $\alpha$  mRNAs have a ~200 nucleotides shorter pA tail than the translational active TNF- $\alpha$  mRNAs.<sup>51</sup> These results suggest that the TNF- $\alpha$  mRNAs present in unstimulated macrophages are translationally silent due to the lack of a pA tail, which represses translation initiation,<sup>52</sup> revealing another mode of regulation of gene expression by polyadenylation.

The IPEX syndrome (immune dysfunctions, polyendocrinopathy, enteropathy, X-linked) is a rare autoimmune disease related to an increase in T-cell activation,<sup>53</sup> due to a mutation in the forkhead box P3 gene (*FOXP3*).<sup>54,55</sup> *FOXP3* 3'UTR analysis in patients revealed an A→G transition in the canonical pA signal after the stop codon (AAUAAA to AAUGAA), which was not present in non-affected controls, which caused a decrease in *FOXP3* expression in affected individuals. This impairment in the polyadenylation and cleavage of *FOXP3* pre-mRNA resulting from a mutation in the pA signal was therefore suggested to be a cause of IPEX.<sup>56</sup>

Similar to IPEX, the Wiskott-Aldrich syndrome (WAS) is an X-linked very rare immunodeficiency<sup>57</sup> related to mutations in the 3'UTR of the Aldrich Syndrome protein (WASP) mRNA. Patients with this syndrome may exhibit multiple phenotypes

and mutations in *WASP* have been described as the principal player in this diversity.<sup>58,59</sup> One of these is a de novo insertion of one adenosine that, together with two deletions, results in the formation of a new pA signal (AAUAAA). As a consequence, *WASP* mRNAs from patients have two alternative transcripts due to APA, instead of only one: the longer transcript is produced by the use of the canonical pA signal and the shortest is made by the use of the newly created pA signal. However, although the two transcripts are translated in two stable mutant proteins, these are non-functional as important domains of the C-terminus are affected.<sup>60</sup>

*Collagen* genes are essential for strength and flexibility of the connective tissue and have a key role in tissue development.<sup>61</sup> The 3'UTR of the *collagen* genes is highly conserved between human and other vertebrates, mainly in the regions surrounding the pA signal, which indicates the presence of important regulatory functions.<sup>62</sup> It has been shown that highly conserved upstream sequence elements (USE) located upstream of *COL1A1*, *COL1A2*, and *COL2A1* pA signal increase polyadenylation efficiency, possibly by providing a binding site for polyadenylation factors.<sup>63</sup> These USEs can consequently be seen as potential agents in controlling the upregulation of *collagen* mRNA in disorders as osteoarthritis and scleroderma.<sup>64,65</sup>

It has recently been shown that Fip-1, a subunit of cleavage and polyadenylation specificity factor (CPSF) that binds to USEs and activates polyadenylation,<sup>66</sup> promotes embryonic stem cell (ESC) self-renewal and somatic cell reprogramming.<sup>67</sup> It was also shown that regulation of APA in pre-mRNAs encoding critical self-renewal factors depend of Fip1-RNA interactions, as well as on the distance between the PAS. When PAS are far apart and Fip1 is highly expressed, as in ESCs, the weaker proximal PAS is usually selected. In contrast, when Fip1 expression is diminished such as in differentiated cells, distal and stronger PAS are used. This may be due to constraints in the binding of CSTF2 to the PAS because when two PAS are positioned close together, the DSE and USE may be juxtaposed or very close together. When Fip-1 levels are high, it binds to the USE and impedes CSTF2 binding to the DSE, which inhibits the usage of the proximal PAS. When Fip1 levels are decreased as in differentiated cells, there is efficient CSTF2 recruitment and utilization of the proximal PAS, producing mRNAs with shorter 3'UTRs. APA regulation by Fip-1 thus seems to serve as a fine-tuning regulatory mechanism in stem cell biology and cell fate specification.<sup>67</sup>

Systemic lupus erythematosus (SLE) is a multisystem autoimmune disease that frequently co-segregates with other immune disorders.<sup>68</sup> Disturbance of regulatory T-cells apoptosis has been related with autoimmunity<sup>69</sup> and possibly contributes to the lymphopenia observed in SLE. As loss of GIMAP5s<sup>70</sup> causes apoptosis of primary T-cells<sup>71</sup> it may be a player in disorders involving lymphopenia and apoptosis. A search for associations between SLE and *GIMAP5* revealed two silent SNPs (single nucleotide polymorphism) in the GIMPAP5 3'UTR, one of them located in the first of three pA signals. The SNP that marks the risk haplotype causes a shift from AAUAAA to AAUAGA, which results in a transcription termination defect, as described for thalassemias. Homozygous individuals for this SNP have higher amounts

of the longer mRNA than heterozygous or homozygous individuals for the canonical pA signal, suggesting that the *GIMAP5* risk haplotype is associated with susceptibility to SLE.<sup>72</sup>

These studies show that alterations in polyadenylation factors (e.g., CSTF2) or in specific pA signals (e.g., in *FOXP3*) cause several immune disorders, highlighting the importance of a tight control of polyadenylation and establishing a possible link between the development of immune responses and APA.

### Neurological diseases

The development of the nervous system is a very complex process and involves a tight network of gene regulation mechanisms. APA that results in longer 3'UTRs has been described in embryonic development and differentiation,<sup>28</sup> in neurons,<sup>73</sup> and in the central nervous system.<sup>74,75</sup> This tissue-specific 3'UTR extension leads to an increase in the complexity of regulation that is particularly significant in the brain, due to the presence of more miRNAs and RBPs binding sites in the mRNA.<sup>76</sup>

The expression levels of Cyclooxygenase 2 (*COX-2*), which is expressed in the brain, are increased during inflammation<sup>77</sup> and regulated by APA.<sup>78,79</sup> *COX-2* produces two different mRNAs (2.8 kb and 4.6 kb) by APA.<sup>80,81</sup> The longest isoform contains a high number of AREs<sup>82</sup> that cause instability and enhance rapid mRNA degradation.<sup>83</sup> It has been demonstrated that the neocortex, which is affected in Alzheimer disease (AD), expresses high levels of 4.6 kb *COX-2* mRNA. Therefore, a possible association between high variability of *COX-2* APA, its expression levels and the heterogeneity of the AD phenotypes may be foreseen.<sup>84</sup>

Mutations in the  $\alpha$ -Synuclein (*aSyn*) gene have been associated with Parkinson disease (PD).<sup>85-87</sup> It has been shown that an *aSyn* mRNA isoform with a longer 3' UTR (*aSynL*) is more highly expressed in brain tissues of PD patients compared with unaffected brains. It has also been shown that a PD risk-associated SNP in the 3'UTR in unaffected individuals is correlated with a higher *aSynL* expression, suggesting that the *aSyn* APA pattern may be used as a biomarker for PD disease. Interestingly, it was shown that *aSynL* mRNA levels are increased by dopamine treatment in the midbrain dopaminergic neurons, indicating that this neurotransmitter modulates *aSyn* APA.<sup>88</sup>

Protein inclusions are the hallmarks of several neurological diseases, including the dominant oculopharyngeal muscular dystrophy (OPMD). OPMD is characterized by ptosis, dysphasia, and proximal limb weakness and filamentous intranuclear inclusions in muscle fibers.<sup>89,90</sup> The mutated gene involved in this disease encodes for poly(A) binding protein, nuclear 1 (PABPN1), also known as *PABP2*. This gene contains a (GCG)<sub>8-13</sub> expansion in dominant OPDM, causing the lengthening of a polyalanine tract located at the N-terminus of the PABPN1 protein.<sup>91</sup> Interestingly, it has been shown that PABPN1 is involved in muscle disease<sup>92,93</sup> and in mouse and human skeletal muscle the steady-state levels of PABPN1 mRNA and protein are lower than in other tissues. However, during muscle regeneration, PABPN1 levels increase, which suggests that PABPN1 is active in muscle repair and explains the nature of OPMD specifically in this tissue. *PABPN1* knockdown in primary mouse myoblasts from extra ocular, pharyngeal and limb muscles causes defects in

myogenesis, particularly in myoblast proliferation and differentiation. It is worth noting that it has also been shown that PABPN1 is required for efficient mRNA export from the nucleus.<sup>93</sup> Intriguingly, mutations in *PABPN1* do not affect the steady-state of the pA tail length. Instead, a high concentration of PABPN1 was found in the nuclear inclusion bodies in the muscle fibers of OPDM patients, whereas in healthy individuals it is dispersed in the nucleoplasm.<sup>94</sup> The mutated protein was more recently shown to function in a dominant-negative manner sequestering the normal protein in nuclear inclusions. This mechanism results in a toxic protein gain-of-function.<sup>95,96</sup> Taking into account the observations that PABPN1 binds to the pA tail of nascent transcripts<sup>97-101</sup> and also that proteins located in the inclusion bodies are generally targeted for degradation it is likely that mutated PABPN1 retains polyadenylated mRNAs in nuclear inclusions. Indeed, mutated PABPN1 affects the cellular localization of polyadenylated mRNAs, altering their function *via* their sequestration in the inclusion bodies, a typical feature of neurological diseases.<sup>94</sup> Interestingly, it was additionally shown that PABPN1 is involved in APA, as its knockdown leads to an increase of proximal cleavage site selection in genes that contain multiple PAS.<sup>96</sup>

Fabry disease is an X-linked rare disorder caused by the deficient activity of lysosomal exoglycosidase,  $\alpha$ -galactosidase A ( $\alpha$ -*GalA*).<sup>102</sup>  $\alpha$ -*GalA* is a peculiar gene because it lacks a 3'UTR and contains the pA signal within the coding sequence.<sup>103</sup> In Fabry disease patients, two frameshift mutations were found in the  $\alpha$ -*GalA* 3' terminus resulting in mutant transcripts with different 3' lengths. The first mutation leads to the transition from AUUAAA to AUUAAG and results in the usage of alternative pA signals or aberrant selection of cleavage sites. The second mutation deletes an ACTT sequence downstream of the pA signal and leads to cleavage of the pre-mRNA at alternative sites. The majority of the mutant  $\alpha$ -*GalA* transcripts were shown to result in non-functional polypeptides causing the development of Fabry disease.<sup>104</sup>

The fragile X mental retardation 1 (*FMRI*) gene produces multiple mRNA isoforms by alternative start sites and APA in Fragile X Syndrome (FXS)-related diseases, such as fragile X-associated immature ovarian insufficiency (FXPOI)<sup>105,106</sup> and fragile X-associated tremor and/or ataxia syndrome (FXTAS).<sup>107,108</sup> Three pA signals were described in the *FMRI* 3'UTR: one is the canonical pA signal and the other two are weaker single nucleotide variants. It was observed that when *FMRI* presents permutation alleles in the 5'UTR (CGG repeats that can be extended from 55 to 200 and do not inactivate the gene<sup>105,106</sup>), the mRNA isoforms derived from the two non-canonical pA variant signals decreased. It was also shown that APA produces mRNAs with different pA tail lengths. Therefore, *FMRI* permutation alleles affect polyadenylation and mRNA production and seem to play a role in pathology.<sup>109</sup>

A variety of studies have shown that U1A, a subunit of U1 spliceosomal small nuclear ribonucleoproteins (snRNPs) complex, which has a main function in splicing also has a role in polyadenylation. In 1998 Gunderson et al. have initially demonstrated that U1snRNP inhibits polyadenylation through

its interaction with PAP at the distal PAS of the bovine papilloma virus (BPV) pre-mRNA. Additionally, another component of the U1snRNP complex, the 70K protein, was identified as a key regulator in this mechanism.<sup>110</sup> More recently, it was shown that the U1A protein inhibits polyadenylation of the Survival Motor Neuron (*SMN*) pre-mRNA.<sup>111</sup> This inhibition is a key process in Spinal Muscular Atrophy (SMA) as this neurodegenerative disease is caused by low levels of SMN,<sup>112</sup> which is implicated in snRNP assembling, namely in U1snRNP biogenesis. Free U1A binds to *SMN* 3'UTR with high affinity and specificity, immediately upstream of the CPSF binding site, inhibiting the cleavage of *SMN* pre-mRNA and resulting in a decrease in the SMN protein production.<sup>111</sup>

A role for U1snRNP was additionally demonstrated in preventing premature transcription termination at PAS localized in introns<sup>113</sup> and this activity has recently been referred to as telescripting.<sup>114</sup> It was shown that a decrease in U1snRNP levels causes preferential usage of proximal pA signals, localized either in the 3'UTR or in intronic regions. Therefore, in addition to its role in splicing, U1 snRNP binds to nascent pre-mRNAs to prevent premature cleavage and polyadenylation in cryptic PAS.<sup>110,113</sup> Interestingly, the U1snRNP levels may be responsible for this event in activated neurons. In neurons activated with forskolin and forskolin/KCl, a shortening of *HOMER1* as well as *Dab1* (genes involved in synaptogenesis) pre-mRNAs was observed. On the other hand, a switch to the production of longer isoforms was observed when U1snRNP was overexpressed.<sup>114</sup> Recently, an extracellular aggregation of U1snRNP components in the neuronal cell bodies of the brains of AD patients that caused defects in RNA processing was described.<sup>115</sup> These results indicate that U1snRNP function in splicing and polyadenylation plays a significant role in AD. Remarkably, low levels of U1snRNP were also observed in a cancer cell line (HeLa)<sup>114</sup> affecting genes that were previously described to undergo 3'UTR shortening in activated T cells.<sup>16</sup> Furthermore, during differentiation and development there is a decrease in 3' end processing activity including U1snRNP,<sup>34,114</sup> and thus it is possible that it contributes to the APA pattern observed in these cellular processes. Taken together, these results demonstrate a key role for U1snRNP in regulating APA events and mRNA expression.

### Endocrine diseases

With the emergence of high-throughput sequencing technologies, multiple polyadenylation alterations responsible for several endocrine diseases were recently revealed. One of these occurs in the steroidogenic acute regulatory (*StAR*) gene, where the encoded protein mediates the crucial step of delivery of cholesterol to the inner mitochondrial membrane in steroidogenic tissues.<sup>116-118</sup> Two *StAR* mRNA transcripts, the short and long isoforms, result from APA due to the usage of two pA signals located in the 3'UTR. These isoforms have the same expression levels in adrenal cells in a basal state.<sup>119</sup> However, upon stimulation with Br-cAMP, which stimulates cholesterol metabolism, there is a preferential production of the longest mRNA isoform due to usage of the distal PAS that is less stable than the shorter

mRNA. Regulation of *StAR* at the mRNA level thus seems to be necessary for rapid regulation of the levels of this important endocrine regulator.<sup>120</sup> Although mutations in *StAR* have already been associated with pathologies such as congenital adrenal hyperplasia (CAH),<sup>121,122</sup> these results further suggest that the impairment of its RNA processing mechanism can also cause alterations in cholesterol metabolism.

Hyperglycemia is the strongest player in Diabetic nephropathy, a fatal complication of type I and type II diabetes mellitus.<sup>123</sup> The high-glucose-regulated gene (*HGRG-14*) was identified as one of the genes differentially expressed in hyperglycemia conditions, where expression undergoes APA regulation in response to high concentrations of glucose. Although in normal conditions cells only express the short transcript (700 bp), after 2 h of hyperglycemic medium incubation they start to express the longest mRNA isoforms (2 Kb). This suggests a lengthening of *HGRG-14* mRNA in hyperglycemic conditions that may be due to a switch in pA signal selection. Interestingly, a rapid *HGRG-14* mRNA decay in high glucose medium was also observed, which may be due to the presence of five AREs in the longer mRNA.<sup>124</sup> Taken together, these results suggest that APA regulates the exclusion (in the short mRNA isoform) or inclusion (in the longer mRNA isoform) of mRNA destabilizing sequences,<sup>124</sup> which mediate mRNA decay.<sup>125</sup> This regulatory mechanism of mRNA stability in hyperglycemic cells provides an example of how hyperglycemia modulates a cellular function via APA.

Another gene that has been consistently related with type II diabetes is the transcription factor 7-like 2 (*TCF7L2*)<sup>126</sup> member of T-cell factor/lymphoid enhancer factor (TCF/LEF) family.<sup>127</sup> Locke *et al.* 2011 identified a pA signal in intron 4 of *TCF7L2* that, when used, produces a truncated mRNA transcript. This truncated transcript is produced in similar amounts to the full-length transcript in human tissues associated with type II diabetes pathogenesis. Similarly to observations made in other studies<sup>128,129</sup> it may be this *TCF7L2* loss of function that causes predisposition to diabetes.<sup>130</sup>

In endocrinological illnesses, as in many other diseases referred in this review, the type of polyadenylation or APA regulation seems to depend of an intrinsic cellular condition.

### Concluding Remarks

This review highlights the relevance of polyadenylation and APA mechanisms in the correct expression of several cellular genes involved in disease development and cell homeostasis. There are definitely many more examples where these pre-mRNA processing mechanisms play crucial roles in disease and in specific cellular conditions. Many studies on polyadenylation and APA have been reported throughout the years and more recently, with the advance of genome wide methodologies, the number of transcripts identified due to APA has increased impressively. However, in some of these cases, the question remains as to whether the shortening in 3'UTR observed for some mRNAs is due to the usage of the proximal PAS or to the



presence of silencing mechanisms such as miRNA operating on the longer mRNA isoforms. As cellular alterations in polyadenylation and/or cleavage factors generally cause dramatic changes in the APA pattern, it is apparent that APA is not controlled by a single master regulator, but by a rather complex mechanism that involves many factors and determinants. Indeed, regulation of polyadenylation relies on the precise integration of the transcription mechanisms with the presence or absence of specific *cis*-elements in the pre-mRNA and *trans*-acting factors in the nucleus, in a particular cellular state. The development of high throughput technologies will undoubtedly lead to a faster identification of polyadenylation and APA associated diseases. Furthermore, it will possibly lead to the identification of disease biomarkers that may be used as diagnostic and therapeutic tools in the future. It is of utmost importance that the molecular mechanisms regulating polyadenylation and APA are elucidated in depth, to understand the development of certain pathologies.

#### Disclosure of Potential Conflicts of Interest

No potential conflict of interest was disclosed.

#### References

1. Lutz CS, Moreira A. Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. *Wiley Interdiscip Rev RNA* 2011; 2:22-31; PMID:21956967; <http://dx.doi.org/10.1002/wrna.47>
2. Tian B, Manley JL. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* 2013; 38:312-20; PMID:23632313; <http://dx.doi.org/10.1016/j.tibs.2013.03.005>
3. Di Giannantonio DC, Nishida K, Manley JL. Mechanisms and consequences of alternative polyadenylation. *Mol Cell* 2011; 43:853-66; PMID:21925375; <http://dx.doi.org/10.1016/j.molcel.2011.08.017>
4. Elkon R, Ugalde AP, Agami R. Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet* 2013; 14:496-506; PMID:23774734; <http://dx.doi.org/10.1038/nrg3482>
5. Shi Y. Alternative polyadenylation: new insights from global analyses. *RNA* 2012; 18:2105-17; PMID:23097429; <http://dx.doi.org/10.1261/rna.035899.112>
6. Chang H, Lim J, Ha M, Kim VN. TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell* 2014; 53:1044-52; PMID:24582499; <http://dx.doi.org/10.1016/j.molcel.2014.02.007>
7. Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* 2014; 508:66-71; PMID:24476825; <http://dx.doi.org/10.1038/nature13007>
8. Jalkanen AL, Coleman SJ, Wilusz J. Determinants and implications of mRNA poly(A) tail size - Does this protein make my tail look big? *Semin Cell Dev Biol* 2014; PMID:24910447; <http://dx.doi.org/10.1016/j.semcdb.2014.05.018>
9. Scorilas A. Polyadenylate polymerase (PAP) and 3' end pre-mRNA processing: function, assays, and association with disease. *Crit Rev Clin Lab Sci* 2002; 39:193-224; PMID:12120781; <http://dx.doi.org/10.1080/10408360290795510>
10. Kondrashov A, Meijer HA, Barthet-Barateig A, Parker HN, Khurshid A, Tessier S, Sicard M, Knox AJ, Pang L, De Moor CH. Inhibition of polyadenylation reduces inflammatory gene induction. *RNA* 2012; 18:2236-50; PMID:23118416; <http://dx.doi.org/10.1261/rna.032391.112>

#### Acknowledgments

We are very grateful to Eugénia Lisboa for critically proof-reading this manuscript. We thank our colleagues and all current and past members of the lab for suggestions and fruitful discussions. We apologize to all colleagues whose work could not be cited due to space constraints.

#### Funding

This work was funded by FEDER through the Operational Competitiveness Programme-COMPETE and by National Funds through FCT (Fundação para a Ciência e a Tecnologia) under the projects COMPETE: FCOMP-01-0124-FEDER-021201 (PTDC/SAU-GMG/116621/2010), FCOMP-01-0124-FEDER-028252 (PTDC/BEX-BCM/0468/2012) and "NORTE-07-0124-FEDER-000003-Cell Homeostasis Tissue Organization and Organism Biology" co-funded by Programa Operacional Regional do Norte (ON.2-O Novo Norte), under the Quadro de Referência Estratégico Nacional (QREN), through the Fundo Europeu de Desenvolvimento Regional (FEDER) and by FCT (Fundação para a Ciência e Tecnologia).

11. Lee C-Y, Chen L. Alternative polyadenylation sites reveal distinct chromatin accessibility and histone modification in human cell lines. *Bioinformatics* 2013; 29:1713-7; PMID:23740743; <http://dx.doi.org/10.1093/bioinformatics/btt288>
12. Pinto PAB, Henriques T, Freitas MO, Martins T, Domingues RG, Wyrzykowska PS, Coelho PA, Carmo AM, Sunkel CE, Proudfoot NJ, et al. RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J* 2011; 30:2431-44; PMID:21602789; <http://dx.doi.org/10.1038/emboj.2011.156>
13. Nagaïke T, Manley JL. Transcriptional activators enhance polyadenylation of mRNA precursors. *RNA Biol* 2011; 8:964-7; PMID:21941122; <http://dx.doi.org/10.4161/rna.8.6.17210>
14. Martincic K, Alkan SA, Cheate A, Borghesi L, Milcarek C. Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. *Nat Immunol* 2009; 10:1102-9; PMID:19749764; <http://dx.doi.org/10.1038/ni.1786>
15. Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* 2009; 138:673-84; PMID:19703394; <http://dx.doi.org/10.1016/j.cell.2009.06.016>
16. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* 2008; 320:1643-7; PMID:18566288; <http://dx.doi.org/10.1126/science.1155390>
17. Orkin SH, Cheng TC, Antonarakis SE, Kazazian HH Jr. Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human beta-globin gene. *EMBO J* 1985; 4:453-6; PMID:4018033
18. Rund D, Dowling C, Najjar K, Rachmilewitz EA, Kazazian HH Jr., Oppenheim A. Two mutations in the beta-globin polyadenylation signal reveal extended transcripts and new RNA polyadenylation sites. *Proc Natl Acad Sci U S A* 1992; 89:4324-8; PMID:1374896; <http://dx.doi.org/10.1073/pnas.89.10.4324>
19. Higgs DR, Goodbourn SEY, Lamb J, Clegg JB, Weatherall DJ, Proudfoot NJ. Alpha-thalassemia caused by a polyadenylation signal mutation. *Nature* 1983; 306:398-400; PMID:6646217; <http://dx.doi.org/10.1038/306398a0>
20. Hartevelde CL, Losekoot M, Haak H, Heister GA, Giordano PC, Bernini LF. A novel polyadenylation signal mutation in the alpha 2-globin gene causing alpha thalassaemia. *Br J Haematol* 1994; 87:139-43; PMID:7947237; <http://dx.doi.org/10.1111/j.1365-2141.1994.tb04883.x>
21. Ferraresi P, Marchetti G, Legnani C, Cavallari E, Castoldi E, Mascoli F, Ardissino D, Palareti G, Bernardi F. The heterozygous 20210 G/A prothrombin genotype is associated with early venous thrombosis in inherited thrombophilias and is not increased in frequency in artery disease. *Arterioscler Thromb Vasc Biol* 1997; 17:2418-22; PMID:9409210; <http://dx.doi.org/10.1161/01.ATV.17.11.2418>
22. Makris M, Preston FE, Beauchamp NJ, Cooper PC, Daly ME, Hampton KK, Bayliss P, Peake IR, Miller GJ. Co-inheritance of the 20210A allele of the prothrombin gene increases the risk of thrombosis in subjects with familial thrombophilia. *Thromb Haemost* 1997; 78:1426-9; PMID:9423788
23. Kyrle PA, Mannhalter C, Béguin S, Stümpflen A, Hirschl M, Weltermann A, Stain M, Brenner B, Speiser W, Pabinger I, et al. Clinical studies and thrombin generation in patients homozygous or heterozygous for the G20210A mutation in the prothrombin gene. *Arterioscler Thromb Vasc Biol* 1998; 18:1287-91; PMID:9714136; <http://dx.doi.org/10.1161/01.ATV.18.8.1287>
24. Gehring NH, Frede U, Neu-Yilik G, Hundsdoerfer P, Vetter B, Hentze MW, Kulozik AE. Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. *Nat Genet* 2001; 28:389-92; PMID:11443298; <http://dx.doi.org/10.1038/ng578>
25. Carter AM, Sachchithanathan M, Stasinopoulos S, Maurer F, Medcalf RL. Prothrombin G20210A is a bifunctional gene polymorphism. *Thromb Haemost* 2002; 87:846-53; PMID:12038788
26. Ceelie H, Spaargaren-van Riel CC, Bertina RM, Vos HL. G20210A is a functional mutation in the prothrombin gene; effect on protein levels and 3'-end formation. *J Thromb Haemost* 2004; 2:119-27; PMID:14717975; <http://dx.doi.org/10.1111/j.1538-7836.2003.00493.x>
27. Leroyer C, Mercier B, Oger E, Chenu E, Abgrall JF, Férec C, Mottier D. Prevalence of 20210 A allele of the prothrombin gene in venous thromboembolism

- patients. *Thromb Haemost* 1998; 80:49-51; PMID:9684784
28. Ji Z, Lee JY, Pan Z, Jiang B, Tian B. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc Natl Acad Sci U S A* 2009; 106:7028-33; PMID:19372383; <http://dx.doi.org/10.1073/pnas.0900028106>
  29. Elkon R, Drost J, van Haafden G, Jenal M, Schrier M, Oude Vrielink JA, Agami R. E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biol* 2012; 13:R59; PMID:22747694; <http://dx.doi.org/10.1186/gb-2012-13-7-r59>
  30. Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR 3<sup>rd</sup>, Frank J, Manley JL. Molecular architecture of the human pre-mRNA 3' processing complex. *Mol Cell* 2009; 33:365-76; PMID:19217410; <http://dx.doi.org/10.1016/j.molcel.2008.12.028>
  31. Fu Y, Sun Y, Li Y, Li J, Rao X, Chen C, Xu A. Differential genome-wide profiling of tandem 3' UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res* 2011; 21:741-7; PMID:21474764; <http://dx.doi.org/10.1101/gr.115-295.110>
  32. Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev* 2013; 27:2380-96; PMID:24145798; <http://dx.doi.org/10.1101/gad.229328.113>
  33. Liaw H-H, Lin C-C, Juan H-F, Huang H-C. Differential microRNA regulation correlates with alternative polyadenylation pattern between breast cancer and normal cells. *PLoS One* 2013; 8:e56958; PMID:23437281; <http://dx.doi.org/10.1371/journal.pone.0056958>
  34. Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods* 2013; 10:133-9; PMID:23241633; <http://dx.doi.org/10.1038/nmeth.2288>
  35. Morris AR, Bos A, Diosdado B, Rooijers K, Elkon R, Bolijn AS, Carvalho B, Meijer GA, Agami R. Alternative cleavage and polyadenylation during colorectal cancer development. *Clin Cancer Res* 2012; 18:5256-66; PMID:22874640; <http://dx.doi.org/10.1158/1078-0432.CCR-12-0543>
  36. Masamha CPXZ, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, Li W, Wagner EJ. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* 2014; 510:412-6; PMID:24814343
  37. Peterson ML. Immunoglobulin heavy chain gene regulation through polyadenylation and splicing competition. *Wiley Interdiscip Rev RNA* 2011; 2:92-105; PMID:21956971; <http://dx.doi.org/10.1002/wrna.36>
  38. Santos P, Arumemi F, Park KS, Borghesi L, Milcarek C. Transcriptional and epigenetic regulation of B cell development. *Immunol Res* 2011; 50:105-12; PMID:21717070; <http://dx.doi.org/10.1007/s12026-011-8225-y>
  39. Early P, Rogers J, Davis M, Calame K, Bond M, Wall R, Hood L. Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. *Cell* 1980; 20:313-9; PMID:6771020; [http://dx.doi.org/10.1016/0092-8674\(80\)90617-0](http://dx.doi.org/10.1016/0092-8674(80)90617-0)
  40. Rogers J, Early P, Carter C, Calame K, Bond M, Hood L, Wall R. Two mRNAs with different 3' ends encode membrane-bound and secreted forms of immunoglobulin mu chain. *Cell* 1980; 20:303-12; PMID:6771019; [http://dx.doi.org/10.1016/0092-8674\(80\)90616-9](http://dx.doi.org/10.1016/0092-8674(80)90616-9)
  41. Takagaki Y, Seipel RL, Peterson ML, Manley JL. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* 1996; 87:941-52; PMID:8945520; [http://dx.doi.org/10.1016/S0092-8674\(00\)82000-0](http://dx.doi.org/10.1016/S0092-8674(00)82000-0)
  42. Phillips C, Jung S, Gunderson SI. Regulation of nuclear poly(A) addition controls the expression of immunoglobulin M secretory mRNA. *EMBO J* 2001; 20:6443-52; PMID:11707415; <http://dx.doi.org/10.1093/emboj/20.22.6443>
  43. Phillips C, Pachikara N, Gunderson SI. U1A inhibits cleavage at the immunoglobulin M heavy-chain secretory poly(A) site by binding between the two downstream GU-rich regions. *Mol Cell Biol* 2004; 24:6162-71; PMID:15226420; <http://dx.doi.org/10.1128/MCB.24.14.6162-6171.2004>
  44. Milcarek C, Albring M, Langer C, Park KS. The eleven-nineteen lysine-rich leukemia gene (ELL2) influences the histone H3 protein modifications accompanying the shift to secretory immunoglobulin heavy chain mRNA production. *J Biol Chem* 2011; 286:33795-803; PMID:21832080; <http://dx.doi.org/10.1074/jbc.M111.272096>
  45. Chuvpilo S, Zimmer M, Kerstan A, Glöckner J, Avots A, Escher C, Fischer C, Inashkina I, Jankevics E, Berberich-Siebelt F, et al. Alternative polyadenylation events contribute to the induction of NF-ATc in effector T cells. *Immunity* 1999; 10:261-9; PMID:10072078; [http://dx.doi.org/10.1016/S1074-7613\(00\)80026-6](http://dx.doi.org/10.1016/S1074-7613(00)80026-6)
  46. Shell SA, Hesse C, Morris SM Jr., Milcarek C. Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. *J Biol Chem* 2005; 280:39950-61; PMID:16207706; <http://dx.doi.org/10.1074/jbc.M508848200>
  47. Beutler B. TNF, immunity and inflammatory disease: lessons of the past decade. *J Investig Med* 1995; 43:227-35; PMID:7614068
  48. Han J, Brown T, Beutler B. Endotoxin-responsive sequences control cachectin/tumor necrosis factor biosynthesis at the translational level. *J Exp Med* 1990; 171:465-75; PMID:2303781; <http://dx.doi.org/10.1084/jem.171.2.465>
  49. Beisang D, Bohjanen PR. Perspectives on the ARE as it turns 25 years old. *Wiley Interdiscip Rev RNA* 2012; 3:719-31; PMID:22733578; <http://dx.doi.org/10.1002/wrna.1125>
  50. Ivanov P, Anderson P. Post-transcriptional regulatory networks in immunity. *Immunol Rev* 2013; 253:253-72; PMID:23550651; <http://dx.doi.org/10.1111/immr.12051>
  51. Ulich TR, Watson LR, Yin SM, Guo KZ, Wang P, Thang H, del Castillo J. The intratracheal administration of endotoxin and cytokines. I. Characterization of LPS-induced IL-1 and TNF mRNA expression and the LPS-, IL-1-, and TNF-induced inflammatory infiltrate. *Am J Pathol* 1991; 138:1485-96; PMID:2053596
  52. Crawford EK, Ensor JE, Kalvakolanu I, Hasday JD. The role of 3' poly(A) tail metabolism in tumor necrosis factor-alpha regulation. *J Biol Chem* 1997; 272:21120-7; PMID:9261116; <http://dx.doi.org/10.1074/jbc.272.34.21120>
  53. Shigeoka AO, Chance PF, Fain P, Barker DA, Book LS, Rallison ML. An X-linked T-cell activation syndrome maps near the Wiskott-Aldrich locus Xp11.2 - diarrhea, respiratory-infections, autoimmune-disease and endocrinopathies in the absence of platelet defects. *Clinical Research* 1993; 41:A41-A.
  54. Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, Kelly TE, Saulsbury FT, Chance PF, Ochs HD. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat Genet* 2001; 27:20-1; PMID:11137993; <http://dx.doi.org/10.1038/83713>
  55. Chatila TA, Blaeser F, Ho N, Lederman HM, Voulgaropoulos C, Helms C, Bowcock AM. JM2, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic dysregulation syndrome. *J Clin Invest* 2000; 106:R75-81; PMID:11120765; <http://dx.doi.org/10.1172/JCI11679>
  56. Bennett CL, Brunkow ME, Ramsdell F, O'Brian KC, Zhu Q, Fuleihan RL, Shigeoka AO, Ochs HD, Chance PF. A rare polyadenylation signal mutation of the FOXP3 gene (AAUAAA->AAUGAA) leads to the IPEX syndrome. *Immunogenetics* 2001; 53:435-9; PMID:11685453; <http://dx.doi.org/10.1007/s002510100358>
  57. Familiärer AW. angeborener Morbus Werlhofii? [Familial, congenital Werlhof's disease?]. *Monatsschr Kinderheilkd* 1937; 68:212-6.
  58. Jin Y, Mazza C, Christie JR, Giliani S, Fiorini M, Mella P, Gandellini F, Stewart DM, Zhu Q, Nelson DL, et al. Mutations of the Wiskott-Aldrich Syndrome Protein (WASP): hotspots, effect on transcription, and translation and phenotype/genotype correlation. *Blood* 2004; 104:4010-9; PMID:15284122; <http://dx.doi.org/10.1182/blood-2003-05-1592>
  59. Snapper SB, Rosen FS. The Wiskott-Aldrich syndrome protein (WASP): roles in signaling and cytoskeletal organization. *Annu Rev Immunol* 1999; 17:905-29; PMID:10358777; <http://dx.doi.org/10.1146/annurev.immunol.17.1.905>
  60. Andreu N, Garcia-Rodríguez M, Volpini V, Frecha C, Molina IJ, Fontan G, Fillat C. A novel Wiskott-Aldrich syndrome protein (WASP) complex mutation identified in a WAS patient results in an aberrant product at the C-terminus from two transcripts with unusual polyA signals. *J Hum Genet* 2006; 51:92-7; PMID:16372137; <http://dx.doi.org/10.1007/s10038-005-0328-7>
  61. Persikov AV, Brodsky B. Unstable molecules form stable tissues. *Proc Natl Acad Sci U S A* 2002; 99:1101-3; PMID:11830649; <http://dx.doi.org/10.1073/pnas.042707899>
  62. Määttä A, Bornstein P, Penttinen RPK. Highly conserved sequences in the 3'-untranslated region of the COL1A1 gene bind cell-specific nuclear proteins. *FEBS Lett* 1991; 279:9-13; PMID:1995349; [http://dx.doi.org/10.1016/0014-5793\(91\)80237-W](http://dx.doi.org/10.1016/0014-5793(91)80237-W)
  63. Natalizio BJ, Muniz LC, Arhin GK, Wilusz J, Lutz CS. Upstream elements present in the 3'-untranslated region of collagen genes influence the processing efficiency of overlapping polyadenylation signals. *J Biol Chem* 2002; 277:42733-40; PMID:12200454; <http://dx.doi.org/10.1074/jbc.M208070200>
  64. Jimenez SA, Hitraya E, Varga J. Pathogenesis of scleroderma. *Collagen. Rheum Dis Clin North Am* 1996; 22:647-74; PMID:8923589; [http://dx.doi.org/10.1016/S0889-857X\(05\)70294-5](http://dx.doi.org/10.1016/S0889-857X(05)70294-5)
  65. Le Graverand MPH, Eggerer J, Vignon E, Otterness IG, Barclay L, Hart DA. Assessment of specific mRNA levels in cartilage regions in a lapine model of osteoarthritis. *J Orthop Res* 2002; 20:535-44; PMID:12038628; [http://dx.doi.org/10.1016/S0736-0266\(01\)00126-7](http://dx.doi.org/10.1016/S0736-0266(01)00126-7)
  66. Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J* 2004; 23:616-26; PMID:14749727; <http://dx.doi.org/10.1038/sj.emboj.7600070>
  67. Lackford B, Yao C, Charles GM, Weng L, Zheng X, Choi E-A, Xie X, Wan J, Xing Y, Freudenberg JM, et al. Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J* 2014; 33:878-89; PMID:24596251; <http://dx.doi.org/10.1002/emboj.201386537>
  68. Nath SK, Kelly JA, Harley JB, Scofield RH. Mapping the systematic lupus erythematosus susceptibility genes. *Methods Mol Med* 2004; 102:11-29; PMID:15286378
  69. Marleau AM, Sarvetnick N. T cell homeostasis in tolerance and immunity. *J Leukoc Biol* 2005; 78:575-84; PMID:15894586; <http://dx.doi.org/10.1189/jlb.0105050>
  70. Stamm O, Krücken J, Schmitt-Wrede HP, Benten WPM, Wunderlich F. Human ortholog to mouse

- gene *imap38* encoding an ER-localizable G-protein belongs to a gene family clustered on chromosome 7q32-36. *Gene* 2002; 282:159-67; PMID:11814688; [http://dx.doi.org/10.1016/S0378-1119\(01\)00837-X](http://dx.doi.org/10.1016/S0378-1119(01)00837-X)
71. MacMurray AJ, Moralejo DH, Kwitek AE, Rutledge EA, Van Yserloo B, Gohlke P, Speros SJ, Snyder B, Schaefer J, Bieg S, et al. Lymphopenia in the BB rat model of type 1 diabetes is due to a mutation in a novel immune-associated nucleotide (lan)-related gene. *Genome Res* 2002; 12:1029-39; PMID:12097339; <http://dx.doi.org/10.1101/gr.412702>
  72. Hellquist A, Zucchelli M, Kivinen K, Saarialho-Kere U, Koskenmies S, Widen E, Julkunen H, Wong A, Karjalainen-Lindsberg ML, Skoog T, et al. The human GIMAP5 gene has a common polyadenylation polymorphism increasing risk to systemic lupus erythematosus. *J Med Genet* 2007; 44:314-21; PMID:17220214; <http://dx.doi.org/10.1136/jmg.2006.046185>
  73. Shepard PJ, Choi E-A, Lu J, Flanagan LA, Hertel KJ, Shi Y. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA* 2011; 17:761-72; PMID:21343387; <http://dx.doi.org/10.1261/rna.2581711>
  74. Hilgers V, Perry MW, Hendrix D, Stark A, Levine M, Haley B. Neural-specific elongation of 3' UTRs during *Drosophila* development. *Proc Natl Acad Sci U S A* 2011; 108:15864-9; PMID:21896737; <http://dx.doi.org/10.1073/pnas.1112672108>
  75. Smibert P, Miura P, Westholm JO, Shenker S, May G, Duff MO, Zhang D, Eads BD, Carlson J, Brown JB, et al. Global patterns of tissue-specific alternative polyadenylation in *Drosophila*. *Cell Rep* 2012; 1:277-89; PMID:22685694; <http://dx.doi.org/10.1016/j.celrep.2012.01.001>
  76. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HTS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 2009; 460:479-86; PMID:19536157
  77. Simon LS. Role and regulation of cyclooxygenase-2 during inflammation. *Am J Med* 1999; 106(5B):37S-42S; PMID:10390126; [http://dx.doi.org/10.1016/S0002-9343\(99\)00115-1](http://dx.doi.org/10.1016/S0002-9343(99)00115-1)
  78. Hall-Pogar T, Zhang H, Tian B, Lutz CS. Alternative polyadenylation of cyclooxygenase-2. *Nucleic Acids Res* 2005; 33:2565-79; PMID:15872218; <http://dx.doi.org/10.1093/nar/gki544>
  79. Hall-Pogar T, Liang S, Hague LK, Lutz CS. Specific trans-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR. *RNA* 2007; 13:1103-15; PMID:17507659; <http://dx.doi.org/10.1261/rna.577707>
  80. Ristimäki A, Narko K, Hla T. Down-regulation of cytokine-induced cyclo-oxygenase-2 transcript isoforms by dexamethasone: evidence for post-transcriptional regulation. *Biochem J* 1996; 318:325-31; PMID:8761489
  81. Ristimäki A, Garfinkel S, Wessendorf J, Maciag T, Hla T. Induction of cyclooxygenase-2 by interleukin-1 alpha. Evidence for post-transcriptional regulation. *J Biol Chem* 1994; 269:11769-75; PMID:8163473
  82. Shaw G, Kamen R. A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* 1986; 46:659-67; PMID:3488815; [http://dx.doi.org/10.1016/0092-8674\(86\)90341-7](http://dx.doi.org/10.1016/0092-8674(86)90341-7)
  83. Hla T, Neilson K. Human cyclooxygenase-2 cDNA. *Proc Natl Acad Sci U S A* 1992; 89:7384-8; PMID:1380156; <http://dx.doi.org/10.1073/pnas.89.16.7384>
  84. Lukiw WJ, Bazan NG. Cyclooxygenase 2 RNA message abundance, stability, and hypervariability in sporadic Alzheimer neocortex. *J Neurosci Res* 1997; 50:937-45; PMID:9452008; [http://dx.doi.org/10.1002/\(SICI\)1097-4547\(19971215\)50:6<937::AID-JNR4>3.0.CO;2-E](http://dx.doi.org/10.1002/(SICI)1097-4547(19971215)50:6<937::AID-JNR4>3.0.CO;2-E)
  85. Spillantini MG, Schmidt ML, Lee VMY, Trojanowski JQ, Jakes R, Goedert M. Alpha-synuclein in Lewy bodies. *Nature* 1997; 388:839-40; PMID:9278044; <http://dx.doi.org/10.1038/42166>
  86. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M, Peuralinna T, Dutra A, Nussbaum R, et al. alpha-Synuclein locus triplication causes Parkinson's disease. *Science* 2003; 302:841; PMID:14593171; <http://dx.doi.org/10.1126/science.1090278>
  87. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 1997; 276:2045-7; PMID:9197268; <http://dx.doi.org/10.1126/science.276.5321.2045>
  88. Rhinn H, Qiang L, Yamashita T, Rhee D, Zolin A, Vanti W, Abeliovich A. Alternative alpha-synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nat Commun* 2012; 3:1084; PMID:23011138; <http://dx.doi.org/10.1038/ncomms2032>
  89. Tomé FMS, Fardeau M. Nuclear inclusions in oculopharyngeal dystrophy. *Acta Neuropathol* 1980; 49:85-7; PMID:6243839; <http://dx.doi.org/10.1007/BF00692226>
  90. Tomé FMS, Chateau D, Helbling-Leclerc A, Fardeau M. Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy. *Neuromuscul Disord* 1997; 7(Suppl 1):S63-9; PMID:9392019; [http://dx.doi.org/10.1016/S0960-8966\(97\)00085-0](http://dx.doi.org/10.1016/S0960-8966(97)00085-0)
  91. Brais B, Bouchard JP, Xie YG, Rochefort DL, Chrétien N, Tomé FMS, Lafrenière RG, Rommens JM, Uyama E, Nohira O, et al. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 1998; 18:164-7; PMID:9462747; <http://dx.doi.org/10.1038/ng0298-164>
  92. Banerjee A, Apponi LH, Pavlath GK, Corbett AH. PABPN1: molecular function and muscle disease. *FEBS J* 2013; 280:4230-50; PMID:23601051; <http://dx.doi.org/10.1111/febs.12294>
  93. Apponi LH, Leung SW, Williams KR, Valentini SR, Corbett AH, Pavlath GK. Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet* 2010; 19:1058-65; PMID:20035013; <http://dx.doi.org/10.1093/hmg/ddp569>
  94. Calado A, Tomé FMS, Brais B, Rouleau GA, Kühn U, Wahle E, Carmo-Fonseca M. Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum Mol Genet* 2000; 9:2321-8; PMID:11001936; <http://dx.doi.org/10.1093/oxfordjournals.hmg.a018924>
  95. Simonelig M. PABPN1 shuts down alternative poly(A) sites. *Cell Res* 2012; 22:1419-21; PMID:22641371; <http://dx.doi.org/10.1038/cr.2012.86>
  96. Jenal M, Elkon R, Loayza-Puch F, van Haften G, Kühn U, Menzies FM, Oude Vrielink JA, Bos AJ, Drost J, Rooijers K, et al. The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* 2012; 149:538-53; PMID:22502866; <http://dx.doi.org/10.1016/j.cell.2012.03.022>
  97. Beelman CA, Parker R. Degradation of mRNA in eukaryotes. *Cell* 1995; 81:179-83; PMID:7736570; [http://dx.doi.org/10.1016/0092-8674\(95\)90326-7](http://dx.doi.org/10.1016/0092-8674(95)90326-7)
  98. Sachs AB, Sarnow P, Hentze MW. Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* 1997; 89:831-8; PMID:9200601; [http://dx.doi.org/10.1016/S0092-8674\(00\)80268-8](http://dx.doi.org/10.1016/S0092-8674(00)80268-8)
  99. Wahle E. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* 1991; 66:759-68; PMID:1878970; [http://dx.doi.org/10.1016/0092-8674\(91\)90119-J](http://dx.doi.org/10.1016/0092-8674(91)90119-J)
  100. Bienroth S, Keller W, Wahle E. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J* 1993; 12:585-94; PMID:8440247
  101. Wahle E. Poly(A) tail length control is caused by termination of processive synthesis. *J Biol Chem* 1995; 270:2800-8; PMID:7852352
  102. Desnick RJ, Allen KY, Desnick SJ, Raman MK, Bernlohr RW, Krivit W. Fabry's disease: enzymatic diagnosis of hemizygotous and heterozygotous. Alpha-galactosidase activities in plasma, serum, urine, and leukocytes. *J Lab Clin Med* 1973; 81:157-71; PMID:4683418
  103. Bishop DF, Kornreich R, Desnick RJ. Structural organization of the human alpha-galactosidase A gene: further evidence for the absence of a 3' untranslated region. *Proc Natl Acad Sci U S A* 1988; 85:3903-7; PMID:2836863; <http://dx.doi.org/10.1073/pnas.85.11.3903>
  104. Yasuda M, Shabbeer J, Osawa M, Desnick RJ. Fabry disease: novel alpha-galactosidase A 3'-terminal mutations result in multiple transcripts due to aberrant 3'-end formation. *Am J Hum Genet* 2003; 73:162-73; PMID:12796853; <http://dx.doi.org/10.1086/376608>
  105. Toniolo D, Rizzolio F. X chromosome and ovarian failure. *Semin Reprod Med* 2007; 25:264-71; PMID:17594607; <http://dx.doi.org/10.1055/s-2007-980220>
  106. Allingham-Hawkins DJ, Babul-Hirji R, Chitayat D, Holden JJ, Yang KT, Lee C, Hudson R, Gorwill H, Nolin SL, Glicksman A, et al. Fragile X premutation is a significant risk factor for premature ovarian failure: the International Collaborative POF in Fragile X study—preliminary data. *Am J Med Genet* 1999; 83:322-5; PMID:10208170; [http://dx.doi.org/10.1002/\(SICI\)1096-8628\(19990402\)83:4<322::AID-AJMG17>3.0.CO;2-B](http://dx.doi.org/10.1002/(SICI)1096-8628(19990402)83:4<322::AID-AJMG17>3.0.CO;2-B)
  107. Hagerman PJ, Hagerman RJ. The fragile-X premutation: a maturing perspective. *Am J Hum Genet* 2004; 74:805-16; PMID:15052536; <http://dx.doi.org/10.1086/386296>
  108. Bourgeois JA, Coffey SM, Rivera SM, Hessel D, Gane LW, Tassone F, Greco C, Finucane B, Nelson L, Berry-Kravis E, et al. A review of fragile X premutation disorders: expanding the psychiatric perspective. *J Clin Psychiatry* 2009; 70:852-62; PMID:19422761; <http://dx.doi.org/10.4088/JCP.08f04476>
  109. Tassone F, De Rubeis S, Carosi C, La Fata G, Serpa G, Raske C, Willemssen R, Hagerman PJ, Bagni C. Differential usage of transcriptional start sites and polyadenylation sites in FMR1 premutation alleles. *Nucleic Acids Res* 2011; 39:6172-85; PMID:21478165; <http://dx.doi.org/10.1093/nar/gkr100>
  110. Gunderson SI, Polycarpou-Schwarz M, Mattaj IW. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell* 1998; 1:255-64; PMID:9659922; [http://dx.doi.org/10.1016/S1097-2765\(00\)80026-X](http://dx.doi.org/10.1016/S1097-2765(00)80026-X)
  111. Workman E, Veith A, Battle DJ. U1A regulates 3' processing of the survival motor neuron mRNA. *J Biol Chem* 2014; 289:3703-12; PMID:24362020; <http://dx.doi.org/10.1074/jbc.M113.538264>
  112. Lefebvre S, Bürglen L, Reboulet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 1995; 80:155-65; PMID:7813012; [http://dx.doi.org/10.1016/0092-8674\(95\)90460-3](http://dx.doi.org/10.1016/0092-8674(95)90460-3)
  113. Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 2010; 468:664-8; PMID:20881964; <http://dx.doi.org/10.1038/nature09479>
  114. Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, et al. U1 snRNP determines mRNA length and regulates isoform expression. *Cell* 2012; 150:53-64; PMID:22770214; <http://dx.doi.org/10.1016/j.cell.2012.05.029>

115. Bai B, Hales CM, Chen P-C, Gozal Y, Dammer EB, Fritz JJ, Wang X, Xia Q, Duong DM, Street C, et al. U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer's disease. *Proc Natl Acad Sci U S A* 2013; 110:16562-7; PMID:24023061; <http://dx.doi.org/10.1073/pnas.1310249110>
116. Krueger RJ, Orme-Johnson NR. Acute adrenocorticotrophic hormone stimulation of adrenal corticosteroidogenesis. Discovery of a rapidly induced protein. *J Biol Chem* 1983; 258:10159-67; PMID:6309771
117. Epstein LF, Orme-Johnson NR. Regulation of steroid hormone biosynthesis. Identification of precursors of a phosphoprotein targeted to the mitochondrion in stimulated rat adrenal cortex cells. *J Biol Chem* 1991; 266:19739-45; PMID:1655794
118. Sugawara T, Holt JA, Driscoll D, Strauss JF 3<sup>rd</sup>, Lin D, Miller WL, Patterson D, Clancy KP, Hart IM, Clark BJ, et al. Human steroidogenic acute regulatory protein: functional activity in COS-1 cells, tissue-specific expression, and mapping of the structural gene to 8p11.2 and a pseudogene to chromosome 13. *Proc Natl Acad Sci U S A* 1995; 92:4778-82; PMID:7761400; <http://dx.doi.org/10.1073/pnas.92.11.4778>
119. Duan H, Cherradi N, Feige J-J, Jefcoate C. cAMP-dependent posttranscriptional regulation of steroidogenic acute regulatory (STAR) protein by the zinc finger protein ZFP36L1/TIS11b. *Mol Endocrinol* 2009; 23:497-509; PMID:19179481; <http://dx.doi.org/10.1210/me.2008-0296>
120. Zhao D, Duan H, Kim YC, Jefcoate CR. Rodent StAR mRNA is substantially regulated by control of mRNA stability through sites in the 3'-untranslated region and through coupling to ongoing transcription. *J Steroid Biochem Mol Biol* 2005; 96:155-73; PMID:16039847; <http://dx.doi.org/10.1016/j.jsbmb.2005.02.011>
121. Lin D, Gitelman SE, Saenger P, Miller WL. Normal genes for the cholesterol side chain cleavage enzyme, P450<sub>scc</sub>, in congenital lipoid adrenal hyperplasia. *J Clin Invest* 1991; 88:1955-62; PMID:1661294; <http://dx.doi.org/10.1172/JCI115520>
122. Tee MK, Lin D, Sugawara T, Holt JA, Guiguen Y, Buckingham B, Strauss JF 3<sup>rd</sup>, Miller WL. T→A transversion 11 bp from a splice acceptor site in the human gene for steroidogenic acute regulatory protein causes congenital lipoid adrenal hyperplasia. *Hum Mol Genet* 1995; 4:2299-305; PMID:8634702; <http://dx.doi.org/10.1093/hmg/4.12.2299>
123. Mauer SM, Steffes MW, Ellis EN, Sutherland DER, Brown DM, Goetz FC. Structural-functional relationships in diabetic nephropathy. *J Clin Invest* 1984; 74:1143-55; PMID:6480821; <http://dx.doi.org/10.1172/JCI111523>
124. Abdel Wahab N, Gibbs J, Mason RM. Regulation of gene expression by alternative polyadenylation and mRNA instability in hyperglycemic mesangial cells. *Biochem J* 1998; 336:405-11; PMID:9820818
125. Wennborg A, Sohlberg B, Angerer D, Klein G, von Gabain A. A human RNase E-like activity that cleaves RNA sequences involved in mRNA stability control. *Proc Natl Acad Sci U S A* 1995; 92:7322-6; PMID:7638189; <http://dx.doi.org/10.1073/pnas.92.16.7322>
126. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R, Nejjari C, Patsch W, Chikri M, Meyre D, et al. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med (Berl)* 2007; 85:777-82; PMID:17476472; <http://dx.doi.org/10.1007/s00109-007-0203-4>
127. Prokunina-Olsson L, Welch C, Hansson O, Adhikari N, Scott LJ, Usher N, Tong M, Sprau A, Swift A, Bonnycastle LL, et al. Tissue-specific alternative splicing of TCF7L2. *Hum Mol Genet* 2009; 18:3795-804; PMID:19602480; <http://dx.doi.org/10.1093/hmg/ddp321>
128. da Silva Xavier G, Loder MK, McDonald A, Tarasov AI, Carzaniga R, Kronenberger K, Barg S, Rutter GA. TCF7L2 regulates late events in insulin secretion from pancreatic islet beta-cells. *Diabetes* 2009; 58:894-905; PMID:19168596; <http://dx.doi.org/10.2337/db08-1187>
129. Shu L, Matveyenko AV, Kerr-Conte J, Cho J-H, McIntosh CHS, Maedler K. Decreased TCF7L2 protein levels in type 2 diabetes mellitus correlate with downregulation of GIP- and GLP-1 receptors and impaired beta-cell function. *Hum Mol Genet* 2009; 18:2388-99; PMID:19386626; <http://dx.doi.org/10.1093/hmg/ddp178>
130. Locke JM, Da Silva Xavier G, Rutter GA, Harries LW. An alternative polyadenylation signal in TCF7L2 generates isoforms that inhibit T cell factor/lymphoid-enhancer factor (TCF/LEF)-dependent target genes. *Diabetologia* 2011; 54:3078-82; PMID:21913056; <http://dx.doi.org/10.1007/s00125-011-2290-6>