

*Review*

## **Impaired Pre-mRNA Processing and Altered Architecture of 3' Untranslated Regions Contribute to the Development of Human Disorders**

**Eva Michalova, Borivoj Vojtesek and Roman Hrstka \***

Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Zluty kopec 7, Brno 656 53, Czech Republic; E-Mails: michalova@mou.cz (E.M.); vojtesek@mou.cz (B.V.)

\* Author to whom correspondence should be addressed; E-Mail: hrstka@mou.cz;  
Tel.: +420-5-4313-3306; Fax: +420-5-4321-1169.

*Received: 31 May 2013; in revised form: 21 June 2013 / Accepted: 24 June 2013 /*

*Published: 26 July 2013*

---

**Abstract:** The biological fate of each mRNA and consequently, the protein to be synthesised, is highly dependent on the nature of the 3' untranslated region. Despite its non-coding character, the 3' UTR may affect the final mRNA stability, the localisation, the export from the nucleus and the translation efficiency. The conserved regulatory sequences within 3' UTRs and the specific elements binding to them enable gene expression control at the posttranscriptional level and all these processes reflect the actual state of the cell including proliferation, differentiation, cellular stress or tumourigenesis. Through this article, we briefly outline how the alterations in the establishment and final architecture of 3' UTRs may contribute to the development of various disorders in humans.

**Keywords:** 3' UTR; trinucleotide repeat expansion; polymorphism; mutation; cancer

---

### **1. Introduction**

According to the central dogma of molecular biology, proteosynthesis proceeds from the genetic information carried by a DNA sequence through its transcription to RNA that functions as a template for polypeptide synthesis during the follow-up translation step [1]. Therefore, RNA might represent an interface between the coding DNA and the final protein. Recently it has also become known that RNAs hold various functions in the cell and not all the RNAs are synthesised as mRNA templates for future polypeptides.

Transcription transfers the genetic information from a coding gene to a primary transcript—A precursor-mRNA (pre-mRNA). Normally, pre-mRNAs undergo a series of posttranscriptional modifications in the nucleus which bring forth the mRNA. These processing steps include capping of the 5' end, removal of introns by splicing, endonucleolytic cleavage and polyadenylation of the 3' end and editing. mRNAs are finally assembled by a central coding region which is translated to the final polypeptide and by non-coding regions at 5' and 3' ends that are not translated (untranslated regions; UTRs).

During the last decade, UTRs have been shown to harbour various sequence motifs (*cis*-acting elements, *cis*-elements) that in cooperation with specific binding proteins or RNAs (*trans*-acting elements, *trans*-elements) regulate the proper posttranscriptional modifications and proteosynthesis. Together with the pre-mRNA processing mechanism, these regulations play important roles in maintaining cellular functions reflecting cellular proliferation or differentiation and participating in cellular reactions to stress stimuli. Therefore, any alteration in the processing steps, mutation or polymorphism can influence the final character of the mRNA 3' end, the functional properties of the 3' UTR and the fate of the mRNA leading to variety of disorders in humans including cancer.

## 2. Processing of Pre-mRNA 3' End

### 2.1. Polyadenylation Signal

The cleavage and the addition of a poly(A) tail to the 3' end of a pre-mRNA are crucial for an efficient transcription termination, mRNA stability and export to the cytoplasm where the polypeptide chain is synthesised on the ribosomes [2,3]. The polyadenylation signal (PAS), a hexamer “AAUAAA” (less frequently—app. 15%—“AUUAAA”), located approximately 10–30 nucleotides (nt) upstream of the cleavage site, was identified as a highly conserved signal for the endonucleolytic cleavage at the 3' end [4]. Although additional sequences are recognised to stimulate the cleavage reaction [5], the PAS remains crucial for this process. Therefore, alterations in PAS sequence (SNP, insertions/deletions) disrupt the cleavage and polyadenylation steps resulting in various pathologies in humans, including an association with different malignancies. For instance, modified hexamers were found to inactivate gene expression in globin coding genes (“AAUAAG” in  $\alpha$ 2-globin, “AACAAA” in  $\beta$ -globin) of thalassaemia patients [6,7]. Similarly, a transition “AAUAAA”→“AAUGAA” within the PAS of *Foxp3* (forkhead box P3) transcript contributes to the IPEX syndrome (Immunodysregulation, Polyendocrinopathy, and Enteropathy, X-linked), a fatal autoimmune disease by reduced levels of Foxp3 transcription factor leading to the dysfunction of regulatory T cells [8]. The polymorphism of PAS in the human *N*-acetyltransferase coding gene (*NAT1*) influences the acetylation of carcinogens and administered drugs. Conversely, the T→A transversion modifying the polyadenylation signal (“AAUAAA”→“AAAAAA”) within the *NAT1\*10* pre-mRNA does not significantly change the final protein level or the catalytic activity as shown previously in bladder and colon tissues [9]. However, the insertion of “AAA” to the 3' side of the PAS in the mutant allele *NAT1\*16* results in a significant decrease of the protein level and of the catalytic activity measured *in vitro* suggesting that the cause lies in the disrupted secondary structure of the mRNA [10]. The human serotonin transporter terminates the neurotransmission by the reuptake of serotonin and the alteration in the transporter

coding gene (*hSERT*) is potentially involved in the development of the affective disorder [11]. Two different polyadenylation sites “AATGAA” and “AG/TTAAC” were identified within *hSERT*, and this polymorphism was considered to be playing role in this event, however it did not correlate with the susceptibility to the affective disorder [12]. Human lysosomal alpha-galactosidase A, the enzyme responsible for glycosphingolipid catabolism, bears the PAS within the coding sequence of the *GAL* gene and its mRNA thus lacks the 3' UTR [13]. The deletion of the “AA” dinucleotide within the PAS results in deficient enzymatic activity (residual or null) of the protein and in the development of Fabry disease, an inborn X-linked disorder characterised by the accumulation of globotriaosylceramide (GL-3), particularly, in vascular endothelial cells throughout the body [14]. Fabry disease patients develop vasculopathy and their life expectancy is shortened due to the renal insufficiency, cardiac disease and stroke [15,16].

## 2.2. Cleavage and Polyadenylation

Numerous proteins grouped into functional protein complexes participate in 3' end processing. The cleavage and polyadenylation specificity factor (CPSF) consists of five subunits: CPSF-160, -100, -73, -30 and hFip1. It recognises the PAS by the CPSF-160 subunit and catalyses the cleavage reaction through the CPSF-73 subunit [2,17]. Cleavage factors I and II recognise the additional sequence elements required for 3' end processing. They provide an interaction with poly(A) polymerase (PAP) and with nuclear poly(A) binding protein (PABPN1) and stimulate the cleavage reaction [18,19]. PAP catalyses the addition of a poly(A) tail under the control of PABPN1 [20].

The pre-mRNA 3' end is cleaved at the pA site, preferentially after the “CA” dinucleotide, however variations were observed, e.g., in the prothrombin gene (coagulation factor II, F2) where the cleavage normally occurs after the “CG” dinucleotide [21]. The cleavage reaction after the “CG” dinucleotide was observed to be less productive *in vitro* [22]. When mutation “CG”→“CA” occurs in F2 mRNA, this more effective 3' end processing leads to an increased concentration of F2 factor in plasma, resulting in a higher risk for thrombosis development [23].

The poly(A) tail comprising approximately 250 A-nucleotides in mammals is attached to the primary transcript at the cleaved pA site by a protein complex with PAP. The poly(A) sequence protects the 3' end against degrading exonucleases and enables the export of mRNA to the cytoplasm. It is also essential for the transcription termination, and together with the 5' cap and related binding proteins it enhances the translation process. The emerging poly(A) sequence is recognised and bound by PABPN1, which needs at least 27 A-nucleotides for a stable protection of the 3' end against nucleases [24]. PABPN1 is an ubiquitously expressed protein which binds to and regulates the processivity of PAP (that normally shows a low affinity for RNA substrates) and thus controls the elongation of poly(A) tail [25]. Normally, 10× “GCN” repeats coding for alanines within the N-terminus of the protein are present at the 5' end of the first exon of *PABPN1* gene. The expansion to 12–17× “GCN” repeats results in the synthesis of a misfolded protein that aggregates as filaments in nuclear inclusions in skeletal muscle fibres leading to cell death [26]. Patients with this expansion of a polyalanine stretch develop oculopharyngeal muscular dystrophy (OPMD), an autosomal dominant muscle disease. OPMD usually occurs at age fifty and manifests by eyelid drooping, mild ophthalmoplegia, dysphagia and generalised muscle weakness and atrophy. Mechanistically, the

deficit of functional PABPN1 affects other pre-mRNA processing steps: polyadenylation, mRNA export to cytoplasm and mRNA stability. Recently PABPN1 was identified to be involved in the regulation of alternative cleavage and polyadenylation (APA) [27]. It is becoming evident that a large portion of human genes contain multiple cleavage sites and PASs in their 3' UTRs generating multiple mRNA isoforms with different 3' UTRs [28–30]. The choice of an alternative pA site determines the length of the 3' UTR and furthermore the stability, localisation and translation efficiency of mRNA [31]. On the basis of the results of multiple studies, the length of 3' UTRs is inversely correlated with mRNA stability, gene expression and cellular proliferation. Mutated PABPN1 strongly correlates with usage of an alternative pA site and therefore deregulated gene expression [32]. An increased level of mutated protein causes cellular stress *in vitro*, enhances the expression of pro-apoptotic proteins and induces apoptosis in a p53 dependent manner [33].

### 3. Repeat Expansion Disorders

Human repeat expansion disorders represent an extensive heterogeneous group of diseases that are caused by the pathological expansion of repeats (mostly trinucleotide) in the coding or non-coding sequences of specific loci. The non-coding repeat expansions cause multisystem diseases and the extent of the repetitions often reflects disease severity and age of onset. The basis of the pathological mechanism of the non-coding repeats expansions thus lies at the posttranscriptional control of the gene expression and the resulting disorders thus mostly show similar aspects. As our review is focused on the role of 3' UTRs, we describe only selected disorders related to this topic.

#### 3.1. Myotonic Dystrophy Type 1

Myotonic dystrophy type 1 (DM1) is an inherited autosomal dominant disease that is caused by the presence of multiple “CUG” repeats within the 3' UTR of myotonic dystrophy protein kinase (*DMPK*) pre-mRNA [34,35]. The number of “CUG” repeats strongly affects the final manifestation of the disease and the age-at-onset and varies from 5 to 35 repeats representing normal alleles, through 50–150 repeats in unaffected individuals or patients with mild to classical syndromes, up to more than 150 repeats in patients with severe DM1 [36]. When more than 1000 repeats are present, foetal development is disrupted and congenital DM1 occurs [37]. Moreover, a higher number of repeats implies an increased inherited instability of the mutant locus and further increase in copies transmitted across generations [38]

“CUG” transcripts were originally thought to cause DM1 pathogenesis when detected in a form of RNA foci in the nuclei of DM1 cells due to a blocked export of mutated transcripts to the cytoplasm [39]. However, a disrupted alternative splicing process in multiple pre-mRNAs was further observed and altered proteins were related to clinical symptoms of DM1 such as insulin resistance, myotonia muscle wasting, cardiac abnormalities or cognitive deficits [40]. The splicing alterations in multiple related transcripts result mainly from an imbalance in the levels of RNA-binding splicing factors CUGBP1 (CUG binding protein 1, also named CUGBP and ETR3-like factor 1, CELF1) and MBNL1 (muscleblind-like 1). There are three types of MBNL in mammals: MBNL1, 2 and 3, with MBNL1 being the best characterised to date. In DM1, CUGBP1 is hyper-phosphorylated by protein kinase C (PKC) and its level increased, whereas MBNL1 and MBNL2 are bound abundantly to

mutated transcripts and retained in the nucleus, co-localised with RNA foci and therefore are not accessible for a correct splicing process in other pre-mRNAs [41,42]. This “sequestration model” for MBNL is strongly supported by the results of the experiments in mouse models where the MBNL1 levels were increased after the transduction and the myotonia was reduced in HSA<sup>LR</sup> poly(CUG) mice (human skeletal beta-actin long repeats = 250 “CUG” repeats) [43]. A similar effect manifested by the splicing aberration was observed *in vitro* in cells transfected with high concentrations of short synthetic oligoribonucleotides composed of “CUG” repeats [44]. The CUGBP1 protein interacts with different pre-mRNAs during muscle development and regulates the processing when phosphorylated by different kinases. The cyclin D3-cdk4/6 complex plays a key role in this process. In DM1 cells, the phosphorylation status of CUGBP1 and its interaction properties are modified due to a decreased level of cyclin D3 [45]. Jones *et al.* demonstrated that cyclin D3 is directed for proteasomal degradation by the GSK3 $\beta$ -mediated phosphorylation when GSK3 $\beta$  (glycogen synthase kinase 3 $\beta$ ) is stabilised by an increased autophosphorylation in DM1 cells induced by the presence of “CUG” transcripts [46]. Additionally, there is an increased and sequestered RNA binding protein hnRNP H, which regulates the alternative splicing leading to aberrant splicing of insulin receptor (INSR) pre-mRNA in DM1 myoblasts [47].

Despite the length of the “CTG” repeat region, its transcription is not blocked and the expanded “CUG” sequences form hairpins or longer dsRNA structures within the transcripts. These RNA secondary structures were shown to be digested *in vitro* with ribonuclease Dicer to shorter (CUG)<sub>n</sub> sequences which may function as endogenous silencers (siRNA) of transcripts containing “CAG” repeats [48]. Moreover, the mutant transcripts interact with several transcription factors (SP1, STAT) disrupting cellular signalling and the transcription of target genes [49].

### 3.2. Huntington's Disease Like 2

The expansion of “CUG” repeats (up to 41 repeats compared to normal cells harbouring 6–28 repeats) within the alternative exon 2a of Junctophilin-3 (*JPH-3*) mRNA leads to the progression of Huntington's Disease Like 2 (HDL-2), an autosomal dominant disorder. The alternative splicing of exon 2a produces transcripts containing repetitions either within the coding region (translated into polyleucine or polyalanine tracts) or within the 3' UTR. HDL-2, similarly to Huntington's disease, manifests by motor defects, neurodegeneration and dementia [50,51]. Nevertheless, HDL-2 patients represent only a minor group of HD-like patients and are mostly Africans or with African remote ancestors [52]. The affected HDL-2 cells are characterised by the presence of RNA foci accumulating the *JPH-3* transcripts with bound MBNL1 resulting in an altered level of JPH-3 protein and in the neurodegeneration in the striatum [53]. Furthermore, a polyglutamine protein translated from antisense “(CAG)<sub>n</sub>” transcripts and accumulated into nuclear inclusions contributes to the pathogenesis via neuronal dysfunction in a mouse model [54]. However, this phenomenon may play a minor role in the progression of HDL-2 and the polyglutamine expanded proteins do not need to be detected in brain samples. Therefore, the sole reduced level of Junctophilin-3, perhaps in association with other cellular processes, might seriously contribute to HDL-2 progression as shown in *JPH-3* knockout mice [55].

#### 4. Altered mRNA Secondary Structure

Mutations within the 3' end might be responsible for the locally altered secondary structures of mRNAs and altered protein characteristics. GATA binding protein 4 (GATA4) is a transcription factor with zinc finger binding motifs that binds to the conserved “GATA” motifs within the promoter sequences of multiple genes [56]. It is essential for myocardial differentiation and function. Mutations found within the coding or non-coding regions lead to serious heart malformations, even death. Different mutations were detected within the 3' UTR and predicted to alter the secondary structure of mRNA potentially affecting transport and localisation of mRNA and reduced binding of GATA4 to target DNA sequences. These reduced transactivations by GATA4 are assumed to directly participate in the congenital heart disease phenotype [56]. Congenital adrenal hyperplasia (CAH) is usually caused by the deficiency of the 21-hydroxylase enzyme encoded by the *CYP21A2* gene. However, in the non-classical form of CAH no mutations in *CYP21A2* were identified, suggesting an involvement of non-coding regulatory regions. Indeed, the \*13 G→A substitutions within the 3' UTR were identified in this subgroup of patients indicating that a substitution within the *CYP21A2* untranslated region is directly involved in the mild form of the disease probably, due to the altered secondary structure of pre-mRNA [57].

Modified regulatory characteristics of 3' UTRs also result from polymorphisms within the specific *cis*-elements or outside them, both leading to the altered secondary structures of mRNAs that affect the accessibility of target binding sites to interacting *trans*-elements [58]. The A→C polymorphism within the miR-155 binding site within the 3' UTR of *AGTR1* (Angiotensin II type 1 receptor) mRNA is responsible for a reduced binding of inhibitory miR-155 leading to an increased level of AGTR1 protein which is associated with hypertension. More than a weaker complementary sequence binding, the altered mRNA secondary structure and a lower accessibility to the miRNA target site may play role in this process [59]. The SNP rs9818870 (C→U polymorphism) in the *MRAS* 3' UTR is associated with a reduced level of M-ras protein in patients with the coronary artery disease (CAD). A higher accessibility of miR-195 to its target site within the *MRAS* 3' UTR and thus higher repression was shown in “U” allele due to a modified mRNA secondary structure even though the SNP is located within the sequence neighbouring the target site [59].

#### 5. Cancer

Tumour development is generally accepted as a multistep process that involves genetic alterations resulting in the gain of oncogene activity and/or the loss of tumour suppressor gene function. Recent reports show that the alterations within 3' UTRs such as mutations leading to the loss of miRNA complementary sites or changes in the length may significantly influence the expression of many genes [60–62]. In addition, many other regulatory sequences (*cis*-elements) within 3' UTRs have been described to date and their alterations or incorrect interplay with specific binding proteins or RNAs (*trans*-elements) are also known to contribute to the malignant phenotype. Here we describe a few examples how variations in the pre-mRNA 3' end processing step are linked with certain aspects of cancer.

Multiple pA sites are present in human pre-mRNA 3' ends and an alternative processing represents another mechanism of the gene expression control [28,63]. The choice of a polyadenylation signal and an alternative cleavage and polyadenylation at more proximal or more distal pA site create mRNA isoforms of different 3' UTR lengths that influence the expression levels mainly through the mRNA stability. Shorter 3' UTR isoforms are generally more stable and produce higher protein levels compared to longer isoforms most likely due to the loss of negative regulatory *cis*-elements and the limited binding sites for miRNAs. Therefore, shorter 3' UTR isoforms are more abundantly synthesised in proliferating and non-differentiated cells including cancer cells [31,64]. By analogy, transformed cells showing a high proliferation rate harbour short 3' UTR isoforms synthesised in highly expressed oncogenes and this correlates with a worse prognosis of the disease [28,61,65].

Three different mutation mechanisms (A insertion at position 1176; 3 base deletion at position 942 and partial duplication of A-rich sequence at position 970) were detected to create a new polyadenylation site within the *CCND1* pre-mRNA. This alternative processing produces shorter and more stable cyclin D1 isoforms in mantle cell lymphoma leading to a higher proliferation rate and shorter survival of patients [66]. The most famous tumour suppressor, p53 protein, and its alterations are frequently discussed in relation to malignant transformation [67]. A SNP "AAUAAA"→"AAUACA" (rs78378222) was identified in the polyadenylation signal of *TP53* pre-mRNA disrupting the pre-mRNA processing and leading to decreased levels of p53 mRNA [68,69]. This transversion was found associated with different cancer types: basal cell carcinoma, prostate cancer, glioma, colorectal adenoma [68], oesophageal squamous cell carcinoma (ESCC) [69], diffuse large B-cell lymphoma (DLBCL) [70] or glioma [71]. Although this proposed mechanism potentially affecting p53 expression can contribute to the malignant phenotype, the role of rs78378222 variant and its prognostic value must be further elucidated. Recently, an alternative polyadenylation of *CDC6* (cell division cycle 6) pre-mRNA was shown to be induced by 17  $\beta$ -oestradiol in oestrogen receptor positive breast cancers cells, revealing new aspects for the posttranscriptional control of gene expression [72].

The poly(A) polymerase (PAP), a member of the multiprotein complex arranging pre-mRNA processing, is mainly responsible for the addition of a poly (A) tail to the cleaved 3' end. This step is crucial in the posttranscriptional regulation of the gene expression. As a higher rate of proteosynthesis is expected in proliferating cells, PAP was also shown to be more active in highly proliferating cells and furthermore more active in cancer cells compared to normal [73]. In this respect, the PAP activity level was shown to reflect the aggressiveness of breast carcinoma or leukaemia and represents an additional prognostic marker in breast carcinoma [74].

The adenylate uridylylate (AU-rich) elements (AREs) are the most common regulatory elements within 3' UTRs influencing the mRNA stability, the translation progress or the alternative RNA processing. AREs mostly destabilise mRNAs and repress the translation through a facilitation of the deadenylation process which results in the accelerated shortening of the poly(A) tail [75]. AREs may be disrupted by a mutation or reduced through the use of alternative PAS in many mRNAs whose elevated expression were related to proto-oncogene activation and a cancer progression. A proto-oncogene, *c-myc*, is constitutively expressed and the *c-myc* mRNA stabilised through a chromosomal translocation disrupting an ARE region in human plasma cell myeloma [76] and Burkitt lymphoma [77]. The enzyme cyclooxygenase 2 (COX-2) encoded by *PTGS2*

(prostaglandin-endoperoxide synthase 2) gene is normally expressed at a low level or even absent because of an efficient regulation at the posttranscriptional level through a repressive ARE within 3' UTR. Normally, ARE directs *PTGS2* mRNA to a rapid decay, and only under a pro-inflammatory or a growth-associated stimuli the COX-2 is rapidly elevated for an increased prostaglandin formation. An enhanced stability of a shorter *PTGS2* mRNA isoform resulting from the alternative cleavage of 3' end and from the loss of repressive AREs was shown to be associated with chronic inflammatory diseases and colorectal cancer [78].

Certain types of cancer are known to be related to cellular transformation by oncogenic viruses. Human papilloma virus type 16 (HPV-16) infection belongs among the factors contributing to cervical cancer development and the viral transforming proteins E6 and E7 play key roles in carcinogenesis [79]. DNA of HPV-16 is often found integrated into the host genome in cancer cells [80,81]. Integrated viral DNAs were shown to express higher levels of E6/E7 mRNAs which were more stable than when expressed from extrachromosomal viral genomes. As the disruption of 3' UTR of early viral region was detected in an integrated HPV-16, a disruption of a potential ARE sequence within this AU-rich region is suggested to participate significantly in this phenomenon [82].

## 6. Future Perspectives

Through this article we aimed to explain the importance of 3' UTRs in the control of gene expression and show a tight relation of the 3' UTR establishment and integrity to a functional proteosynthesis and a healthy cell. As the age of "RNA research" is still at the beginning, many questions about the roles of mRNA sequence/structure relationships and related RNA-binding proteins in the regulation processes remain to be elucidated.

Formerly, the exploration of mechanisms regulating gene expression was mainly oriented towards promoter regions and to related binding transcription factors and the 3' untranslated regions, and their possible regulatory role at the posttranscriptional level were rather omitted. However, as many conserved sequences representing specific regulatory elements were discovered within the 3' end, the 3' UTRs were brought into focus during the last two decades. Different gene expression regulatory mechanisms related to 3' UTRs were described to date, taking effect during cellular proliferation, differentiation, stress conditions and carcinogenesis as well. The disruption of the posttranscriptional events including effector protein complexes mostly represents a primary cause of 3' UTR alteration, touching every step of the processing machinery. Secondly, mutations of regulatory sequences within 3' UTRs and in proteins or ncRNAs (non-coding RNAs) binding to these sequence also occur and lead to a range of disorders in humans. The spectrum of 3' UTR related disorders is wide and based on the results of recent studies may occur across all human diseases.

## Acknowledgements

This work was supported by GACR P206/12/G151, GACR P301/11/1678, MH CZ-DRO (MMCI, 00209805) and by the European Regional Development Fund and the State Budget of the Czech Republic (RECAMO, CZ.1.05/2.1.00/03.0101).



## Conflict of Interest

The authors declare no conflict of interest.

## References

1. Crick, F. Central dogma of molecular biology. *Nature* **1970**, *227*, 561–563.
2. Zhao, J.; Hyman, L.; Moore, C. Formation of mRNA 3' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 405–445.
3. Lemay, J.F.; Lemieux, C.; St-Andre, O.; Bachand, F. Crossing the borders: Poly(a)-binding proteins working on both sides of the fence. *RNA Biol.* **2010**, *7*, 291–295.
4. Proudfoot, N.J.; Brownlee, G.G. 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* **1976**, *263*, 211–214.
5. Matoulkova, E.; Michalova, E.; Vojtesek, B.; Hrstka, R. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol.* **2012**, *9*, 563–576.
6. Higgs, D.R.; Goodbourn, S.E.; Lamb, J.; Clegg, J.B.; Weatherall, D.J.; Proudfoot, N.J. Alpha-thalassaemia caused by a polyadenylation signal mutation. *Nature* **1983**, *306*, 398–400.
7. Serjeant, G.R.; Serjeant, B.E.; Fraser, R.A.; Hambleton, I.R.; Higgs, D.R.; Kulozik, A.E.; Donaldson, A. Hb s-beta-thalassemia: Molecular, hematological and clinical comparisons. *Hemoglobin* **2011**, *35*, 1–12.
8. Bennett, C.L.; Brunkow, M.E.; Ramsdell, F.; O'Briant, K.C.; Zhu, Q.; Fuleihan, R.L.; Shigeoka, A.O.; Ochs, H.D.; Chance, P.F. A rare polyadenylation signal mutation of the foxp3 gene (aauaaa→aaugaa) leads to the ipex syndrome. *Immunogenetics* **2001**, *53*, 435–439.
9. Bell, D.A.; Badawi, A.F.; Lang, N.P.; Ilett, K.F.; Kadlubar, F.F.; Hirvonen, A. Polymorphism in the n-acetyltransferase 1 (nat1) polyadenylation signal: Association of nat1\*10 allele with higher n-acetylation activity in bladder and colon tissue. *Cancer Res.* **1995**, *55*, 5226–5229.
10. de Leon, J.H.; Vatsis, K.P.; Weber, W.W. Characterization of naturally occurring and recombinant human n-acetyltransferase variants encoded by nat1. *Mol. Pharmacol.* **2000**, *58*, 288–299.
11. Harmar, A.J.; Ogilvie, A.D.; Battersby, S.; Smith, C.A.; Blackwood, D.H.; Muir, W.J.; Fink, G.; Goodwin, G.M. The serotonin transporter gene and affective disorder. *Cold Spring Harb. Symp. Quant. Biol.* **1996**, *61*, 791–795.
12. Battersby, S.; Ogilvie, A.D.; Blackwood, D.H.; Shen, S.; Muqit, M.M.; Muir, W.J.; Teague, P.; Goodwin, G.M.; Harmar, A.J. Presence of multiple functional polyadenylation signals and a single nucleotide polymorphism in the 3' untranslated region of the human serotonin transporter gene. *J. Neurochem.* **1999**, *72*, 1384–1388.
13. Bishop, D.F.; Kornreich, R.; Desnick, R.J. Structural organization of the human alpha-galactosidase a gene: Further evidence for the absence of a 3' untranslated region. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 3903–3907.

14. Brady, R.O.; Gal, A.E.; Bradley, R.M.; Martensson, E.; Warshaw, A.L.; Laster, L. Enzymatic defect in fabry's disease. Ceramidetrihexosidase deficiency. *N. Engl. J. Med.* **1967**, *276*, 1163–1167.
15. MacDermot, K.D.; Holmes, A.; Miners, A.H. Natural history of fabry disease in affected males and obligate carrier females. *J. Inherit Metab. Dis.* **2001**, *24*, 13–14.
16. Vedder, A.C.; Linthorst, G.E.; van Breemen, M.J.; Groener, J.E.; Bemelman, F.J.; Strijland, A.; Mannens, M.M.; Aerts, J.M.; Hollak, C.E. The dutch fabry cohort: Diversity of clinical manifestations and gb3 levels. *J. Inherit Metab. Dis.* **2007**, *30*, 68–78.
17. Tian, B.; Manley, J.L. Alternative cleavage and polyadenylation: The long and short of it. *Trends Biochem. Sci.* **2013**, *38*, 312–320.
18. Brown, K.M.; Gilmartin, G.M. A mechanism for the regulation of pre-mrna 3' processing by human cleavage factor im. *Mol. Cell* **2003**, *12*, 1467–1476.
19. de Vries, H.; Ruegsegger, U.; Hubner, W.; Friedlein, A.; Langen, H.; Keller, W. Human pre-mrna cleavage factor ii(m) contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J.* **2000**, *19*, 5895–5904.
20. Bienroth, S.; Keller, W.; Wahle, E. Assembly of a processive messenger rna polyadenylation complex. *The EMBO J.* **1993**, *12*, 585–594.
21. Gehring, N.H.; Frede, U.; Neu-Yilik, G.; Hundsdoerfer, P.; Vetter, B.; Hentze, M.W.; Kulozik, A.E. Increased efficiency of mrna 3' end formation: A new genetic mechanism contributing to hereditary thrombophilia. *Nature Genet.* **2001**, *28*, 389–392.
22. Chen, F.; MacDonald, C.C.; Wilusz, J. Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res.* **1995**, *23*, 2614–2620.
23. Poort, S.R.; Rosendaal, F.R.; Reitsma, P.H.; Bertina, R.M. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* **1996**, *88*, 3698–3703.
24. Steinman, R.A. Mrna stability control: A clandestine force in normal and malignant hematopoiesis. *Leukemia* **2007**, *21*, 1158–1171.
25. Kerwitz, Y.; Kuhn, U.; Lilie, H.; Knoth, A.; Scheuermann, T.; Friedrich, H.; Schwarz, E.; Wahle, E. Stimulation of poly(a) polymerase through a direct interaction with the nuclear poly(a) binding protein allosterically regulated by ma. *EMBO J.* **2003**, *22*, 3705–3714.
26. Brais, B.; Bouchard, J.P.; Xie, Y.G.; Rochefort, D.L.; Chretien, N.; Tome, F.M.; Lafreniere, R.G.; Rommens, J.M.; Uyama, E.; Nohira, O.; *et al.* Short gcg expansions in the pabp2 gene cause oculopharyngeal muscular dystrophy. *Nature Genet.* **1998**, *18*, 164–167.
27. Jenal, M.; Elkon, R.; Loayza-Puch, F.; van Haaften, G.; Kuhn, U.; Menzies, F.M.; Oude Vrielink, J.A.; Bos, A.J.; Drost, J.; Rooijers, K.; *et al.* The poly(a)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **2012**, *149*, 538–553.
28. 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*, 741–747.
29. Ozsolak, F.; Kapranov, P.; Foissac, S.; Kim, S.W.; Fishilevich, E.; Monaghan, A.P.; John, B.; Milos, P.M. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* **2010**, *143*, 1018–1029.

30. Shepard, P.J.; Choi, E.A.; Lu, J.; Flanagan, L.A.; Hertel, K.J.; Shi, Y. Complex and dynamic landscape of rna polyadenylation revealed by pas-seq. *RNA* **2011**, *17*, 761–772.
31. Ji, Z.; Lee, J.Y.; 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. USA* **2009**, *106*, 7028–7033.
32. de Klerk, E.; Venema, A.; Anvar, S.Y.; Goeman, J.J.; Hu, O.; Trollet, C.; Dickson, G.; den Dunnen, J.T.; van der Maarel, S.M.; Raz, V.; *et al.* Poly(a) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucleic Acids Res.* **2012**, *40*, 9089–9101.
33. Bhattacharjee, R.B.; Zannat, T.; Bag, J. Expression of the polyalanine expansion mutant of nuclear poly(a)-binding protein induces apoptosis via the p53 pathway. *Cell Biol. Int.* **2012**, *36*, 697–704.
34. Fu, Y.H.; Pizzuti, A.; Fenwick, R.G., Jr.; King, J.; Rajnarayan, S.; Dunne, P.W.; Dubel, J.; Nasser, G.A.; Ashizawa, T.; de Jong, P.; *et al.* An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **1992**, *255*, 1256–1258.
35. Mahadevan, M.; Tsilfidis, C.; Sabourin, L.; Shutler, G.; Amemiya, C.; Jansen, G.; Neville, C.; Narang, M.; Barcelo, J.; O'Hoy, K.; *et al.* Myotonic dystrophy mutation: An unstable ctg repeat in the 3' untranslated region of the gene. *Science* **1992**, *255*, 1253–1255.
36. Kamsteeg, E.J.; Kress, W.; Catalli, C.; Hertz, J.M.; Witsch-Baumgartner, M.; Buckley, M.F.; van Engelen, B.G.; Schwartz, M.; Scheffer, H. Best practice guidelines and recommendations on the molecular diagnosis of myotonic dystrophy types 1 and 2. *Eur. J. Hum. Genet.* **2012**, *20*, 1203–1208.
37. Salehi, L.B.; Bonifazi, E.; Stasio, E.D.; Gennarelli, M.; Botta, A.; Vallo, L.; Iraci, R.; Massa, R.; Antonini, G.; Angelini, C.; *et al.* Risk prediction for clinical phenotype in myotonic dystrophy type 1: Data from 2,650 patients. *Genet. Test.* **2007**, *11*, 84–90.
38. Savic Pavicevic, D.; Miladinovic, J.; Brkusanin, M.; Svikovic, S.; Djurica, S.; Brajuskovic, G.; Romac, S. Molecular genetics and genetic testing in myotonic dystrophy type 1. *BioMed Res. Int.* **2013**, doi:10.1155/2013/391821.
39. Taneja, K.L.; McCurrach, M.; Schalling, M.; Housman, D.; Singer, R.H. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell Biol.* **1995**, *128*, 995–1002.
40. Day, J.W.; Ranum, L.P. Rna pathogenesis of the myotonic dystrophies. *Neuromuscul. Disord.* **2005**, *15*, 5–16.
41. Jiang, H.; Mankodi, A.; Swanson, M.S.; Moxley, R.T.; Thornton, C.A. Myotonic dystrophy type 1 is associated with nuclear foci of mutant rna, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* **2004**, *13*, 3079–3088.
42. Kuyumcu-Martinez, N.M.; Wang, G.S.; Cooper, T.A. Increased steady-state levels of cugbp1 in myotonic dystrophy 1 are due to pkc-mediated hyperphosphorylation. *Mol. Cell* **2007**, *28*, 68–78.
43. Kanadia, R.N.; Shin, J.; Yuan, Y.; Beattie, S.G.; Wheeler, T.M.; Thornton, C.A.; Swanson, M.S. Reversal of rna missplicing and myotonia after muscleblind overexpression in a mouse poly(cug) model for myotonic dystrophy. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11748–11753.

44. Mykowska, A.; Sobczak, K.; Wojciechowska, M.; Kozlowski, P.; Krzyzosiak, W.J. Cag repeats mimic cug repeats in the misregulation of alternative splicing. *Nucleic Acids Res.* **2011**, *39*, 8938–8951.
45. Salisbury, E.; Sakai, K.; Schoser, B.; Huichalaf, C.; Schneider-Gold, C.; Nguyen, H.; Wang, G.L.; Albrecht, J.H.; Timchenko, L.T. Ectopic expression of cyclin d3 corrects differentiation of dm1 myoblasts through activation of rna cug-binding protein, cugbp1. *Exp. Cell Res.* **2008**, *314*, 2266–2278.
46. Jones, K.; Wei, C.; Iakova, P.; Bugiardini, E.; Schneider-Gold, C.; Meola, G.; Woodgett, J.; Killian, J.; Timchenko, N.A.; Timchenko, L.T. Gsk3beta mediates muscle pathology in myotonic dystrophy. *J. Clin. Investig.* **2012**, *122*, 4461–4472.
47. Paul, S.; Dansithong, W.; Kim, D.; Rossi, J.; Webster, N.J.; Comai, L.; Reddy, S. Interaction of muscleblind, cug-bp1 and hnnp h proteins in dm1-associated aberrant ir splicing. *EMBO J.* **2006**, *25*, 4271–4283.
48. Krol, J.; Fiszer, A.; Mykowska, A.; Sobczak, K.; de Mezer, M.; Krzyzosiak, W.J. Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. *Mol. Cell* **2007**, *25*, 575–586.
49. Ebralidze, A.; Wang, Y.; Petkova, V.; Ebralidse, K.; Junghans, R.P. Rna leaching of transcription factors disrupts transcription in myotonic dystrophy. *Science* **2004**, *303*, 383–387.
50. Greenstein, P.E.; Vonsattel, J.P.; Margolis, R.L.; Joseph, J.T. Huntington's disease like-2 neuropathology. *Mov. Disord.* **2007**, *22*, 1416–1423.
51. Margolis, R.L.; O'Hearn, E.; Rosenblatt, A.; Willour, V.; Holmes, S.E.; Franz, M.L.; Callahan, C.; Hwang, H.S.; Troncoso, J.C.; Ross, C.A. A disorder similar to huntington's disease is associated with a novel cag repeat expansion. *Ann. Neurol.* **2001**, *50*, 373–380.
52. Paradisi, I.; Ikonomu, V.; Arias, S. Huntington disease-like 2 (hdl2) in venezuela: Frequency and ethnic origin. *J. Hum. Genet.* **2013**, *58*, 3–6.
53. Rudnicki, D.D.; Holmes, S.E.; Lin, M.W.; Thornton, C.A.; Ross, C.A.; Margolis, R.L. Huntington's disease-like 2 is associated with cug repeat-containing rna foci. *Ann. Neurol.* **2007**, *61*, 272–282.
54. Wilburn, B.; Rudnicki, D.D.; Zhao, J.; Weitz, T.M.; Cheng, Y.; Gu, X.; Greiner, E.; Park, C.S.; Wang, N.; Sopher, B.L.; *et al.* An antisense cag repeat transcript at jph3 locus mediates expanded polyglutamine protein toxicity in huntington's disease-like 2 mice. *Neuron* **2011**, *70*, 427–440.
55. Seixas, A.I.; Holmes, S.E.; Takeshima, H.; Pavlovich, A.; Sachs, N.; Pruitt, J.L.; Silveira, I.; Ross, C.A.; Margolis, R.L.; Rudnicki, D.D. Loss of junctophilin-3 contributes to huntington disease-like 2 pathogenesis. *Ann. Neurol.* **2012**, *71*, 245–257.
56. Reamon-Buettner, S.M.; Cho, S.H.; Borlak, J. Mutations in the 3'-untranslated region of gata4 as molecular hotspots for congenital heart disease (chd). *BMC Med. Genet.* **2007**, *8*, 38.
57. Menabo, S.; Balsamo, A.; Baldazzi, L.; Barbaro, M.; Nicoletti, A.; Conti, V.; Pirazzoli, P.; Wedell, A.; Cicognani, A. A sequence variation in 3' utr of cyp21a2 gene correlates with a mild form of congenital adrenal hyperplasia. *J. Endocrinol. Investig.* **2012**, *35*, 298–305.
58. Abrahams, Y.; Laguette, M.J.; Prince, S.; Collins, M. Polymorphisms within the col5a1 3'-utr that alters mrna structure and the mir608 gene are associated with achilles tendinopathy. *Ann. Hum. Genet.* **2013**, *77*, 204–214.

59. Haas, U.; Sczakiel, G.; Laufer, S.D. MicroRNA-mediated regulation of gene expression is affected by disease-associated snps within the 3'-utr via altered rna structure. *RNA Biol.* **2012**, *9*, 924–937.
60. Lee, Y.S.; Dutta, A. The tumor suppressor microRNA let-7 represses the hmga2 oncogene. *Genes Dev.* **2007**, *21*, 1025–1030.
61. Mayr, C.; Bartel, D.P. Widespread shortening of 3' utrs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **2009**, *138*, 673–684.
62. Mayr, C.; Hemann, M.T.; Bartel, D.P. Disrupting the pairing between let-7 and hmga2 enhances oncogenic transformation. *Science* **2007**, *315*, 1576–1579.
63. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; *et al.* The transcriptional landscape of the mammalian genome. *Science* **2005**, *309*, 1559–1563.
64. Sandberg, R.; Neilson, J.R.; Sarma, A.; Sharp, P.A.; Burge, C.B. Proliferating cells express mrnas with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **2008**, *320*, 1643–1647.
65. Lembo, A.; Di Cunto, F.; Provero, P. Shortening of 3'utrs correlates with poor prognosis in breast and lung cancer. *PLoS One* **2012**, doi:10.1371/journal.pone.0031129.
66. Wiestner, A.; Tehrani, M.; Chiorazzi, M.; Wright, G.; Gibellini, F.; Nakayama, K.; Liu, H.; Rosenwald, A.; Muller-Hermelink, H.K.; Ott, G.; *et al.* Point mutations and genomic deletions in ccnd1 create stable truncated cyclin d1 mrnas that are associated with increased proliferation rate and shorter survival. *Blood* **2007**, *109*, 4599–4606.
67. Hrstka, R.; Coates, P.J.; Vojtesek, B. Polymorphisms in p53 and the p53 pathway: Roles in cancer susceptibility and response to treatment. *J. Cell. Mol. Med.* **2009**, *13*, 440–453.
68. Stacey, S.N.; Sulem, P.; Jonasdottir, A.; Masson, G.; Gudmundsson, J.; Gudbjartsson, D.F.; Magnusson, O.T.; Gudjonsson, S.A.; Sigurgeirsson, B.; Thorisdottir, K.; *et al.* A germline variant in the tp53 polyadenylation signal confers cancer susceptibility. *Nature Genet.* **2011**, *43*, 1098–1103.
69. Zhou, L.; Yuan, Q.; Yang, M. A functional germline variant in the p53 polyadenylation signal and risk of esophageal squamous cell carcinoma. *Gene* **2012**, *506*, 295–297.
70. Li, Y.; Gordon, M.W.; Xu-Monette, Z.Y.; Visco, C.; Tzankov, A.; Zou, D.; Qiu, L.; Montes-Moreno, S.; Dybkaer, K.; Orazi, A.; *et al.* Single nucleotide variation in the tp53 3' untranslated region in diffuse large b-cell lymphoma treated with rituximab-chop: A report from the international dlbcl rituximab-chop consortium program. *Blood* **2013**, *121*, 4529–4540.
71. Enciso-Mora, V.; Hosking, F.J.; Di Stefano, A.L.; Zelenika, D.; Shete, S.; Broderick, P.; Idbaih, A.; Delattre, J.Y.; Hoang-Xuan, K.; Marie, Y.; *et al.* Low penetrance susceptibility to glioma is caused by the tp53 variant rs78378222. *Br. J. Cancer* **2013**, *108*, 2178–2185.
72. Akman, B.H.; Can, T.; Erson-Bensan, A.E. Estrogen-induced upregulation and 3'-utr shortening of cdc6. *Nucleic Acids Res.* **2012**, *40*, 10679–10688.
73. Kazazoglou, T.; Tsiapalis, C.M.; Havredaki, M. Polyadenylate polymerase activity in stationary and growing cell cultures. *Exp. Cell Biol.* **1987**, *55*, 164–172.
74. Scorilas, A.; Talieri, M.; Ardavanis, A.; Courtis, N.; Dimitriadis, E.; Yotis, J.; Tsiapalis, C.M.; Trangas, T. Polyadenylate polymerase enzymatic activity in mammary tumor cytosols: A new independent prognostic marker in primary breast cancer. *Cancer Res.* **2000**, *60*, 5427–5433.

75. Xu, N.; Chen, C.Y.; Shyu, A.B. Modulation of the fate of cytoplasmic mRNA by AU-rich elements: Key sequence features controlling mRNA deadenylation and decay. *Mol. Cell. Biol.* **1997**, *17*, 4611–4621.
76. Hollis, G.F.; Gazdar, A.F.; Bertness, V.; Kirsch, I.R. Complex translocation disrupts c-myc regulation in a human plasma cell myeloma. *Mol. Cell. Biol.* **1988**, *8*, 124–129.
77. Eick, D.; Piechaczyk, M.; Henglein, B.; Blanchard, J.M.; Traub, B.; Kofler, E.; Wiest, S.; Lenoir, G.M.; Bornkamm, G.W. Aberrant c-myc mRNAs of Burkitt's lymphoma cells have longer half-lives. *EMBO J.* **1985**, *4*, 3717–3725.
78. Young, L.E.; Dixon, D.A. Posttranscriptional regulation of cyclooxygenase 2 expression in colorectal cancer. *Curr. Colorectal Cancer Rep.* **2010**, *6*, 60–67.
79. zur Hausen, H. Human papillomaviruses and their possible role in squamous cell carcinomas. *Curr. Top. Microbiol. Immunol.* **1977**, *78*, 1–30.
80. Schwarz, E.; Freese, U.K.; Gissmann, L.; Mayer, W.; Roggenbuck, B.; Stremlau, A.; zur Hausen, H. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature* **1985**, *314*, 111–114.
81. Yee, C.; Krishnan-Hewlett, I.; Baker, C.C.; Schlegel, R.; Howley, P.M. Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. *Am. J. Pathol.* **1985**, *119*, 361–366.
82. Jeon, S.; Lambert, P.F. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of e6 and e7 mRNAs: Implications for cervical carcinogenesis. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 1654–1658.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).