BioFormosa（2012）47: 1-7

Studies of Cap Independent Internal Ribosome Entry Segment Activity of *ATXN8OS* RNA

Hsin-Chieh Shiau[1], I-Cheng Chen[1], Hsuan-Yuan Lin, Ming-Tsan Su, Guey-Jen Lee-Chen*
Department of Life Science, National Taiwan Normal University
Taipei, Taiwan
[1]These authors contributed equally to this work

(Received: 25 July 2012, accepted: 10 September 2012)

ABSTRACT

Spinocerebellar ataxia type 8 (SCA8) involves bidirectional expression of CUG (*ATXN8OS*) and CAG (*ATXN8*) expansion transcripts. Although being transcribed, alternatively spliced and polyadenylated in the CTG orientation, only small open reading frames (ORFs) were noted in *ATXN8OS* RNA. To investigate if a 102 amino acids containing-ORF (ORF1) and a 41 amino acids plus a polyleucine tract containing-ORF (ORF2) in the *ATXN8OS* RNA could be translated, we cloned 5' region of the human *ATXN8OS* cDNA into dicistronic *Renilla* and firefly luciferase reporter. Transient transfection in human embryonic kidney (HEK)-293 cells and reporter assay revealed the presence of bipartite cap independent internal ribosome entry segments (IRESs) in the *ATXN8OS* RNA. The translation of ORF1 was further demonstrated by expression of chimeric construct with an in-frame ORF1-EGFP gene. Western blot with GFP antibody suggested that the expressed ORF1-EGFP protein may be initiated from an upstream non-AUG initiator. The biological meaning of the *ATXN8OS* ORF1 protein and its role in the pathogenesis of SCA8 remains to be determined.

Key words: *ATXN8OS*, ORF, IRES, non-AUG initiator

Introduction

Spinocerebellar ataxias (SCAs) comprise a heterogeneous group of disorders involving progressive degeneration of the cerebellum, brainstem, and spinal tract (Wullner, 2003). Among the SCAs, SCA type 8 (SCA8) was described as a CTG repeat expansion in the *ATXN8OS* (ataxin 8 opposite strand) gene and a CAG repeat expansion in the overlapping *ATXN8* (ataxin 8) gene (Ikeda *et al.*, 2008). In the CTG direction, *ATXN8OS* expresses spliced and polyadenylated untranslated transcripts in various brain tissues (Koob *et al.*, 1999). In the CAG direction, the expanded *ATXN8* encodes a polyglutamine expansion protein (Moseley *et al.*, 2006) known to be pathogenic in other polyglutamine disorders. The plausible mechanisms related to the transcripts in the CTG direction include a RNA gain-of function mechanism (Mutsuddi *et al.*, 2004) and an anti-sense RNA interference mechanism (He *et al.*, 2006). Our study using an *ATXN8OS* cellular model also revealed that SCA8 larger triplet expansion alters histone modification and induces RNA foci (Chen *et al.*, 2009).

Although being apparently non-coding (Koob *et al.*, 1999), a 102 amino acids containing-open reading frame 1 (ORF1) 446 nucleotides from the 5' end of *ATXN8OS* RNA and a 41 amino acids plus a polyleucine tract containing-open reading frame 2 (ORF2) 1051 nucleotides from the 5' end of *ATXN8OS* RNA exist (Fig. 1A). The two ORFs could be translated if *ATXN8OS* RNA possesses a cap independent internal ribosome entry site (IRES) activity, which requires the formation of a complex RNA structural element and the presence of *trans*-acting factors (Stoneley and Willis, 2004). In the present study we first examined the cap independent IRES activity in the *ATXN8OS* RNA using a dual luciferase reporter assay. Then the translation of ORF1 was investigated by transient expression of EGFP-tagged ORF1 in conjunction with Western blot and fluorescent microscopy examination.

Materials and Methods

*Corresponding author: Guey-Jen Lee-Chen; FAX: 886-2-29312904; E-mail: t43019@ntnu.edu.tw

Dual luciferase reporter constructs

ATXN8OS cDNA containing exons D, C2, C1, B, and A (Nemes *et al.*, 2000) (Fig. 1A) was cloned as described (Chen *et al.*, 2009). The *ATXN8OS* cDNA were then cloned into the *Eco*RI site of pEGFP-N1 (Clontech). To construct a dual luciferase reporter, a 76-bp *Xba*I-*Bam*HI polylinker region of pcDNA3 was first added between the *Xba*I and *Bam*HI sites of phRL-TK vector (Promega) to introduce a *Xho*I site as well as remove the SV40 late poly(A) region. Then a 1972-bp *Xho*I-*Bam*HI fragment containing the firefly luciferase gene and the SV40 late poly(A) signal from pGL3-Basic vector (Promega) was placed between the *Xho*I and *Bam*HI sites of the modified phRL-TK vector. The resulting pRF plasmid had *Renilla* luciferase and firefly luciferase genes between the TK promoter and polyadenylation signal (Fig. 1B).

To engineer dual luciferase reporter constructs containing different 5' regions of human *ATXN8OS*, the *ATXN8OS* cDNA in pEGFP-N1 was restricted with *Xho*I, *Sty*I, and/or *Hae*III to generate various 5' portion of *ATXN8OS* cDNA (NR_002717; 1~1032, 1~395 and 396~1032) (Fig. 1A). The blunted *Xho*I-*Sty*I, *Xho*I-*Hae*III and *Hae*III-*Sty*I fragments were placed in the blunted *Xho*I site between the two luciferase genes. The 632-bp blunted *Xho*I-*Msc*I IRES fragment from pIRES2-EGFP (Clontech) was inserted between the two luciferase genes as a positive control.

Cell culture, transfections and enzyme assay

Human embryonic kidney (HEK-293) cells cultivated in Dulbecco's modified Eagle's medium containing 10% FCS were plated into 12-well dishes (2×10^5/well), grown for 20 hr and transfected by the lipofection method (GibcoBRL) with the test dual luciferase reporter plasmid (1.5 μg). The cells were grown for 48 hr. Then cell lysates were prepared and luciferase activity was measured by a luminometer using a dual luciferase assay system (Promega). The IRES activity of each *ATXN8OS* cDNA fragment was directly measured by the ratio of the firefly luciferase level to the *Renilla* luciferase level. For each construct, three independent transfection experiments were performed.

ORF1-EGFP construct

The ORF translation termination sequence in C1 exon was removed and a *Sma*I restricted site (underlined) added by PCR using primer 5'-GCG CCCGGGACACTTCAACTTCCTATACATACA. The *Eco*RI (in MCS of pGEM-T Easy vector)-*Sma*I fragment containing *ATXN8OS* ORF was in-frame fused to the EGFP gene in the pEGFP-N1 vector (between the *Eco*RI and *Bst*UI sites). Portion of the Kozak consensus translation initiation sequence (ACCATG) in the EGFP gene was further removed by site-directed mutagenesis (primer 5'-CGGGCCC GGGATCCACCGGTCGCCΔGTGAGCAAGGGC GAGGAGCTG, Δ=ACCATG). The resulting pORF1-EGFP construct (Fig. 2A) was verified by DNA sequencing.

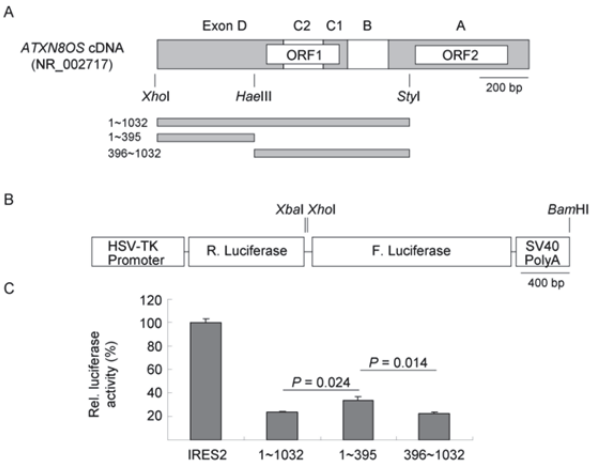

Figure 1. IRES activity of the *ATXN8OS* transcript. (A) The *ATXN8OS* cDNA (NR_002717) contains exons D, C2, C1, B, and A. The putative ORF1 and ORF2 are indicated by the open boxes inside the cDNA. The restriction enzymes and the cutting sites used to generate 5' fragments of *ATXN8OS* are shown on the bottom of the cDNA. (B) The dual luciferase reporter plasmid had *Renilla* luciferase and firefly luciferase genes between the TK promoter and polyadenylation signal. The locations of *Xba*I, *Xho*I and *Bam*HI sites used for construction are shown on the top. (C) Relative luciferase activities generated by dual luciferase constructs with ECMV IRES and various *ATXN8OS* cDNA fragments (1~1032, 1~395, 396~1032) in HEK-293 cells. Forty-eight hours following transfection, cells were harvested and luciferases activities measured. IRES activity is expressed as percentages of the activity of the ECMV IRES, which was set at 100%. Each value is the mean ± SD of three independent experiments each performed in duplicate.

Western blot analysis

HEK-293 cells (6×10^5/6-well) were transfected with the pORF1-EGFP or pIRES2-

EGFP plasmid (4 µg) and grown for 48 hours. Cells were lysed in RIPA buffer (10 mM Tris pH 7.5, 150 mM NaCl, 5 mM EDTA pH 8.0, 1% sodium deoxycholate, 1% NP-40 and 0.1% SDS) containing the protease inhibitor mixture (Sigma). After sonication and sitting on ice for 20 min, the lysates were centrifuged at 14,000 rpm for 30 min at 4°C. Protein concentrations were determined with the Bio-Rad protein assay kit, using albumin as standards. Proteins (25 µg) were electrophoresed on 10% SDS-polyacrylamide gel and transferred onto nitrocellulose membrane (Schleicher and Schuell) by reverse electrophoresis. After blocking, the membrane was stained with anti-GFP (1:200 dilution, Santa Cruz Biotechnology) antibody. The immune complexes were detected using horseradish peroxidase-conjugated goat anti-rabbit (Rochland) IgG antibody (1:10000 dilution) and Immobilon™ Western Chemiluminescent HRP substrate (Millipore).

Confocal microscopy examination

HEK-293 cells (2×10^5/12-well) on coverslips were transfected with the pORF1-EGFP or pIRES2-EGFP plasmid (2 µg) and grown for 48 hours. For visualizing intracellular ORF1-EGFP or EGFP protein, transfected cells were fixed in 4% paraformaldehyde for 15 min. Nuclei were detected using 0.05% DAPI (4'-6-diamidino-2-phenylindole). Cells were examined after mounted in Vectashield (Vector Laboratories Inc.) for GFP and DAPI fluorescence using a Leica TCS confocal laser scanning microscope optimized for simultaneous dual fluorescent imaging.

Results

IRES activity of ATXN8OS RNA

As described, two small ORFs in the *ATXN8OS* transcripts were noted (Fig. 1A). To investigate if these two ORFs can be translated via a cap independent IRES activity, we constructed a dicistronic vector pRF in which firefly luciferase was placed after the *Renilla* luciferase (Fig. 1B). The expression construct was under the control of the HSV-TK promoter. Regions upstream of *ATXN8OS* ORF1 and ORF2 (Fig. 1A) were inserted into the intercistronic region of the pRF. The IRES from the encephalomyocarditis virus (ECMV) (Gurtu *et al.*, 1996) was inserted as a positive control. When the expressed luciferase level of the

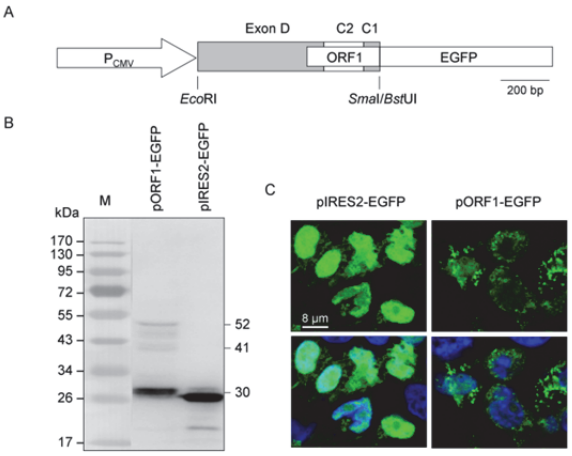

Figure 2. Expression of *ATXN8OS* ORF1-EGFP construct in HEK-293 cells. (A) A 752-bp cDNA fragment containing exons D, C2 and portion of C1 was inserted into pEGFP-N1 multiple cloning sites so that *ATXN8OS* ORF1 was fused in-frame with the EGFP gene to generate pORF1-EGFP. (B) Western blot analysis of cells transfected with indicated plasmids using GFP antibody. (C) Confocal images of cells expressing EGFP and ORF1-EGFP (green). Nuclei were counterstained with DAPI (blue).

ECMV IRES was set as 100%, the *ATXN8OS* 5' fragments 1~1032, 1~395 and 396~1032 directed firefly luciferase synthesis to a level of 23.5%, 33.7% and 22.5%, respectively, as compared to the ECMV IRES sequences (Fig. 1C). The results suggest the possible IRES activity existing in the 5' regions upstream of *ATXN8OS* ORF1 and ORF2.

ATXN8OS ORF1 expression

As the *ATXN8OS* fragment 1~395 expressed significant higher level of relative luciferase activity as compared to fragment 1~1032 (33.7% vs. 23.5%, $P = 0.024$) and fragment 396~1032 (33.7% vs. 22.5%, $P = 0.014$), the translation of ORF1 was further investigated. An EGFP tag was in-frame fused to the C terminal of the *ATXN8OS* ORF1 to generate pORF1-EGFP construct (Fig. 2A). The construct was transiently expressed into HEK-293 cells for two days for protein and EGFP fluorescence studies. On Western blot immunostaining with GFP antibody, 52 kDa, 41 kDa and 30 kDa proteins were detected in pORF1-EGFP transfected cells, as compared to a 26 kDa protein in pIRES2-EGFP transfected cells (Fig. 2B). The observed 41 kDa protein is likely to be the predicted ORF1-EGFP protein (350 amino acids with MW of 39677). Confocal microscopy

examination of GFP fluorescence was carried out to examine the expression of ORF-EGFP protein. As shown in Fig. 2C, strong GFP fluorescence was distributed diffusely in pIRES2-EGFP-transfected cells. With pORF1-EGFP construct, small and dispersed granules appeared mainly in the cytoplasm, in addition to showing diffuse cytoplasm expression.

Discussion

SCA8 was first proposed to be caused by an RNA gain-of-function mechanism and analysis of *ATXN8OS* sequence did not reveal any possible spliced isoform possessing an ORF to extend through the expansion (Ranum and Day, 2004). To study the plausible pathogenesis of SCA8, we cloned the *ATXN8OS* cDNA containing spliced exons D, C2, C1, B, and A (Nemes *et al.*, 2000). Sequence analysis revealed the existence of a 102 amino-acid ORF1 and a 41 amino acids plus a polyleucine tract ORF2 in *ATXN8OS* RNA (Fig. 1A). To investigate if these two ORFs can be translated via a cap independent IRES activity, we cloned regions upstream of *ATXN8OS* ORF1 and ORF2 between firefly and *Renilla* luciferase genes. By comparing the expression levels of firefly luciferase and *Renilla* luciferase among these constructs after transfection, we could possibly define the IRES activity of *ATXN8OS*. As both non-overlapping 1~395 and 396~1032 fragments displayed IRES activity, the presence of multiple internal ribosome entry sites was indicated (Fig. 1). Interestingly, the lower luciferase activity of fragment 1-1032 was observed when compared with fragment 1-395. This result may be due to the specialized structure of different RNA fragment which could recruit ribosome or protein factors. Proteins or factors implicated in the regulation of IRES activity are worthy of further investigation.

To further examine the translation of ORF1, ORF1-EGFP construct (Fig. 2A) was transiently expressed into HEK-293 cells for two days. In addition to the predicted 41 kDa ORF1-EGFP protein, immunoblot with anti-GFP antibody suggested alternative translation initiation from an upstream non-AUG and an in-frame downstream AUG codon, resulting in the production of 52 and 30 kDa isoforms, respectively (Fig. 2B). Fluorescence microscope examination of the expressed GFP-tagged ORF1 protein also provided important information that *ATXN8OS* ORF1 is indeed translatable and may be localized to organelles inside the cytoplasm (Fig. 2C). The candidate IRES sequences for translation of these proteins need to be further determined.

Recently, three yeast genes, CARP2A (acidic ribosomal protein P2A in *C. albicans*) (Abramczyk *et al.*, 2003), GRS1 (one of the two glycyl-tRNA synthetase genes in *S. cerevisiae*) (Chang and Wang, 2004), and ALA1 (AlaRS gene in *S. cerevisiae*) (Tang *et al.*, 2004) have been reported to use naturally occurring non-AUG triplets as translation initiators. In the case of CARP2A, a non-AUG codon GUG serves as the exclusive translation initiator, while in the case of ALA1 and GRS1, non-AUG codons ACG and UUG act as alternative translation initiators that are accompanies by a downstream in-frame AUG initiation codon. The expressed *ATXN8OS* ORF1 may use an upstream in-frame GUG triplet as a native translation initiation codon (Fig. 3). As alternative initiation sites are utilized for the synthesis of proteins that regulate biological processes in health and disease (Lock *et al.*, 1991; Hann *et al.*, 1994; Bruening and Pelletier, 1996), the biological meaning of the *ATXN8OS* ORF1 protein and its role in the pathogenesis of SCA8 remains to be determined.

In conclusion, as the findings in this study indicated, the *ATXN8OS* putative ORF1 protein could be expressed via a naturally occurring non-AUG start codon. Further studies are warranted to reveal of the role of *ATXN8OS* ORF1 and its connection to SCA8.

Acknowledgements

We thank the Molecular Imaging Core Facility of National Taiwan Normal University under the auspices of the National Science Council. This work was supported by grants NSC95-3112-B-003-002 and NSC97-2311-B-003-010-MY3 from the National Science Council, Executive Yuan, Taiwan.

References

Abramczyk D, Tchorzewski M and Grankowski N. 2003. Non-AUG translation initiation of mRNA encoding acidic ribosomal P2A protein in *Candida albicans*. Yeast 20:1045-1052.

Bruening W and Pelletier J. 1996. A non-AUG

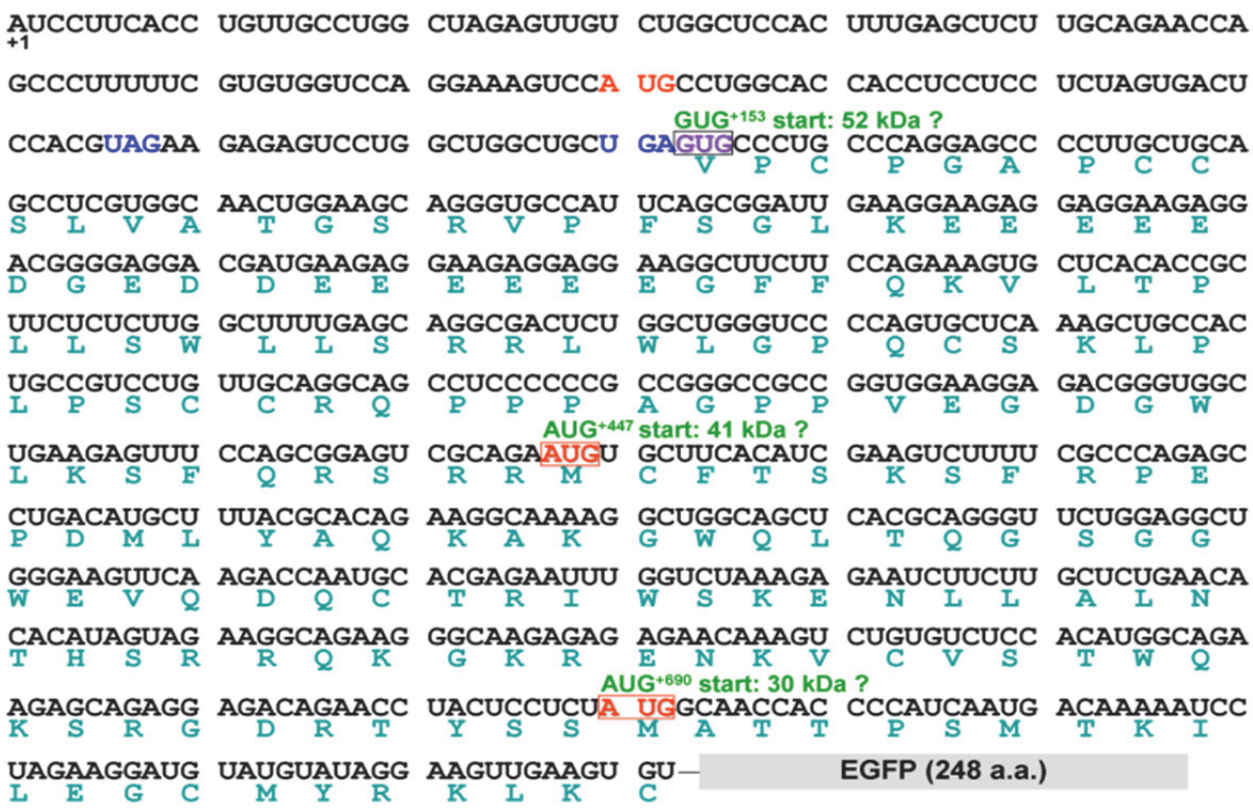

Figure 3. The sequence of ATXN8OS RNA (+1 ~ +752) with GUG^{+153} (marked with purple), AUG^{+447} and AUG^{+690} (marked with red) indicated. Also shown are the predicted amino acid sequences initiated from GUG^{+153} (marked with green).

translational initiation event generates novel WT1 isoforms. J. Biol. Chem. 271:8646-8654.

Chang KJ and Wang CC. 2004. Translation initiation from a naturally occurring non-AUG codon in *Saccharomyces cerevisiae*. J. Biol. Chem. 279:13778-13785.

Chen IC, Lin HY, Lee GC, Kao SH, Chen CM, Wu YR, Hsieh-Li HM, Su MT and Lee-Chen G.J. 2009. Spinocerebellar ataxia type 8 larger triplet expansion induces compact chromatin and RNA foci. BMC Mol. Biol. 10:9.

Gurtu V, Yan G and Zhang G. 1996. IRES bicistronic expression vectors for efficient creation of stable mammalian cell lines. Biochem. Biophys. Res. Commun. 229:295-298.

Hann SR, Dixit M, Sears RC and Sealy L. 1994. The alternatively initiated c-Myc proteins differentially regulate transcription through a noncanonical DNA-binding site. Genes Dev. 8:2441-2452.

He Y, Zu T, Benzow KA, Orr HT, Clark HB and Koob MD. 2006. Targeted deletion of a single Sca8 ataxia locus allele in mice causes abnormal gait, progressive loss of motor coordination, and Purkinje cell dendritic deficits. J. Neurosci. 26:9975-9982.

Hellen CU and Sarnow P. 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. 15:1593-1612.

Ikeda Y, Daughters RS and Ranum LP. 2008. Bidirectional expression of the SCA8 expansion mutation: One mutation, two genes. Cerebellum 7:150-158.

Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW and Ranum LP. 1999. An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). Nat. Genet. 21:379-384.

Lock P, Ralph S, Stanley E, Boulet I, Ramsay R and Dunn AR. 1991. Two isoforms of murine hck, generated by utilization of alternative translational initiation codons, exhibit different patterns of subcellular localization. Mol. Cell. Biol. 11:4363-4370.

Moseley ML, Zu T, Ikeda Y, Gao W, Mosemiller

AK, Daughters RS, Chen G, Weatherspoon MR, Clark HB, Ebner TJ, Day JW and Ranum LP. 2006. Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. Nat. Genet. 38:758-769.

Mutsuddi M, Marshall CM, Benzow KA, Koob MD and Rebay I. 2004. The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in *Drosophila*. Curr. Biol. 14:302-308.

Nemes JP, Benzow KA, Moseley ML, Ranum LP and Koob MD. 2000. The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). Hum. Mol. Genet. 9:1543-1551.

Ranum LP and Day JW. 2004. Pathogenic RNA repeats: an expanding role in genetic disease. Trends Genet. 20:506-512.

Stoneley M and Willis AE. 2004. Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. Oncogene 23:3200-3207.

Tang HL, Yeh LS, Chen NK, Ripmaster T, Schimmel P and Wang CC. 2004. Translation of a yeast mitochondrial tRNA synthetase initiated at redundant non-AUG codons. J. Biol. Chem. 279:49656-49663.

Wullner U. 2003. Genes implicated in the pathogenesis of spinocerebellar ataxias. Drugs Today 39:927-937.

生物學報（2012）47：1-7

ATXN8OS RNA 上不依賴 Cap 的內部核醣體進入位置活性研究

蕭欣杰[1] 陳怡辰[1] 林玄原 蘇銘燦 李桂楨*
國立臺灣師範大學生命科學系
[1]此二作者對本文的貢獻相同

（收稿日期：2012.7.25，接受日期：2012.9.10）

摘　要

第八型脊髓小腦運動失調症與*ATXN8OS* CUG重複擴增RNA及*ATXN8*多麩醯胺擴增的蛋白相關。雖然在CTG方向被轉錄、選擇性裁接及加上polyA尾，*ATXN8OS* RNA上僅存在小的開放解讀架構。為探究*ATXN8OS* RNA上一包含102胺基酸之ORF1及一包含41胺基酸與多白胺酸之ORF2是否被轉譯，我們選殖了人的*ATXN8OS* cDNA 5’片段，置入水母及螢火蟲雙冷光酶基因間。短暫轉染人類胚胎腎細胞及冷光報告測驗顯示，*ATXN8OS* RNA上帶有兩個不依賴Cap的內部核醣體進入位置。進一步表現在同一解讀架構上的ORF1-EGFP基因亦證實ORF1的轉譯。結合GFP抗體的西方轉漬分析顯示，上游的non-AUG密碼可能作為ORF1-EGFP蛋白的轉譯起始密碼。*ATXN8OS* ORF1的生理意義及其在SCA8致病機轉的角色仍有待釐清。

關鍵詞：*ATXN8OS*、開放解讀架構、內部核醣體進入位置、非AUG轉譯起始密碼

*通信作者：李桂楨（Guey-Jen Lee-Chen）；FAX：886-2-29312904；E-mail：t43019@ntnu.edu.tw