行政院國家科學委員會專題研究計畫　期中進度報告

3’-端非轉譯區內三聯核酸重複序列之長度變化對線蟲神經肌肉功能影響之研究(2/3)

計畫類別：個別型計畫
計畫編號：NSC94-2320-B-040-012-
執行期間：94 年 08 月 01 日至 95 年 07 月 31 日
執行單位：中山醫學大學生物醫學科學學系

計畫主持人：蕭光明

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 95 年 5 月 26 日

行政院國家科學委員會補助專題研究計畫 □ 成 果 報 告 □期中進度報告

3’-端非轉譯區內三聯核酸重複序列之長度變化對線蟲神經肌肉功能影響之研究（2/3）

計畫類別：V 個別型計畫　　□ 整合型計畫
計畫編號：NSC 94－2320－B－040－012－
執行期間：　94 年　8 月　1 日至　95 年　7 月　31 日

計畫主持人：蕭光明
共同主持人：
計畫參與人員：潘惠錦、王麗鈞

成果報告類型(依經費核定清單規定繳交)：V 精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：中山醫學大學 生物醫學科學學系

中　華　民　國　95 年　5 月　26 日

一、中文摘要

含 CUG 及 CAG 序列之 RNA 會形成髮夾結構而且擴增之後會在核內聚集成焦。然而，到目前為止只有擴增之 CUG 序列在基因轉殖小鼠體內產生病理效應。在果蠅動物體無論是 CUG 或 CAG 序列擴增之後皆不會有毒性發生。為了釐清含重複序列之 RNA 在病理上的角色，我們利用線蟲作為模式動物將含不轉譯之 CTG 及 CAG 三聯核酸重複序列之轉殖基因表達在蟲體體壁肌肉細胞。結果顯示這二種重複序列會造成與序列長度相關之病理效應。表達 120 次重複序列之線蟲其咽喉肌電圖、肌肉協調性、壽命、及後代數目皆明顯異常。而且其肌肉細胞形態隨個體生長而出現變異。表達 200 次重複序列之線蟲大部分在胚胎時期即死亡，少數孵出之幼蟲則出現生長遲滯現象，其體長停留在 L2 階段無法產生下一代。這些發現顯示重複序列之 CUG RNA 即足以造成個體發育之缺陷。此外我們的結果強烈除了 CTG 序列之外，過長不轉譯的 CAG 序列也會經 RNA 的機制對生物體造成損害。

關鍵詞：基因轉殖線蟲，CTG/CAG 重複序列，RNA 病理機制，神經肌肉結構功能分析

Abstract

Transcripts containing expanded CUG or CAG repeats form hairpin structure *in vitro* and nuclear foci *in vivo*. However, only expanded CUG repeats were shown to mediate pathogenic effects in transgenic *mice*. Both large CUG and CAG repeats are not toxic to *Drosophila*. To clarify the pathological role of the repeat RNA, we have here exploited the nematode *Caenorhabditis elegans* as a model system by expressing transgenes containing various lengths of untranslated CTG and CAG repeats in body wall muscles. Our data demonstrate that both CUG and CAG repeats are deleterious to this simple animal with a threshold of repeat size between 83 and 120. Animals expressing 120 repeats, but not those expressing 83 or shorter repeats, show short life span, small brood size, deviated electrophysiological response, uncoordinated motility, and an age-dependent abnormality in muscle structure. Expression of over 200 repeats results in embryonic lethality or retarded growth at larva stages. These findings suggest that the repeat CUG RNA alone is sufficient to cause developmental defects in a length-dependent manner. Besides, they first demonstrate that, in addition to the CUG repeats, the CAG repeat expansions can also be pathogenic at RNA level.

Key words: Transgenic *C. elegans*, CTG/CAG repeats, RNA pathogenic mechanism, neuromuscular function analysis.

二、緣由與目的

Myotonic dystrophy 1 (DM1) is a dominantly inherited neuromuscular disorder caused by an expansion of CTG repeats in the 3’-untranslated region (UTR) of the dystrophia myotonica protein kinase (*DMPK*) gene. In normal population, the repeat number ranges in size from 5-37. Mild DM1 patients typically have fewer than 100 repeats and only present with cataracts. Classical DM1 patients usually have more than 100 repeats with an onset age around adolescence and develop multiple symptoms, including myotonia, progressive muscle weakness and atrophy, cardiac conduction defects, insulin intolerance, infertility, and reduced life expectancy. Congenitally affected children typically have more than 1000 CTG repeats and are characterized by hypotonia and mental retardation. In addition to DM1 mutation, patients with an CCTG repeat expansion in intron 1 of zinc-finger protein 9 (*ZNF9*) gene (DM2) also develop many typical clinical manifestations of DM1 with the exception that they do not show a congenital form.

The findings that both transcribed but not translated CTG and CCTG repeats causing

similar phenotypes indicate a toxic gain of function of the expanded repeat RNA. Consistently, mice expressing expanded CUG repeats in an unrelated transcript develop typical features of DM phenotypes. The current evidences suggest that the RNAs transcribed from the expanded alleles form hairpin structure and cause *trans*-dominant effects on splicing by altering the functions of CUG RNA binding proteins, including the CUG-BP1/ETR-3 like factors (CELF) and the muscleblind-like (MBNL) proteins. Both CELF and MBNL proteins regulate alternative splicing with antagonistic effects on a subset of genes, such as the *chloride channel 1* (*ClC1*) and *insulin receptor* (*IR*).The steady levels of CUG-BP1 and its binding activity are increased in DM patients. Transgenic mice expressing high levels of CUG-BP1 show neonatal lethality, skeletal muscle histological abnormalities, and aberrant regulation of alternative splicing as observed in congenital DM1. MBNL proteins colocalize with CUG and CCUG RNA foci. Targeted deletion of *Mbnl1* homolog in mice, although does not cause neonatal lethality, results in myopathy, myotonia and splicing defects as seen in classical DM.

The orthologs of *CELF* and *MBNL* genes have been evolutionarily conserved from *C. elegans*, *Drosophila* to *human*, indicating the functional importance of these RNA-binding proteins in multicellular organisms. Indeed, the decreased expression of ETR-1, a homologue of CUG-BP, causes severe defect in muscle development in *C. elegans*. The MBNL homolog in *Drosophila* is required for terminal muscle differentiation. Therefore, it is very likely that the expanded CUG repeats are deleterious to lower eukaryotes through affecting the functions of these RNA-binding proteins as well. Nevertheless, *Drosophila* expressing $(CUG)_{162}$ repeats do not develop any detectable pathological phenotype. The full length human *SCA8* transcripts containing an untranslated $(CUG)_{112}$ repeats tract cause *Drosophila* neurodegeneration. However, *Drosophila* expressing *SCA8* transcripts with $(CUG)_9$ repeats also develop similar phenotype. These observations indicate that the CUG RNA was not toxic to *Drosophila*. The nematode *C. elegans* has been used as a model system in a number of triplet repeat diseases, especially to address the underlying pathogenic mechanisms of the polyQ diseases. To investigate the pathological effects of untranslated triplet repeats in lower eukaryotes, we have, in this study, explored *C. elegans* as a model system by expressing CTG repeats in the 3'-UTR of a reporter gene. Because CAG RNA also form hairpin structure and nuclear foci which colocalize with MBNL, we have studied the effects of CAG repeats RNA in parallel.

三、結果與討論

Generation of transgenic *C. elegans*

To study the pathogenic effects of untranslated CTG repeats on *C. elegans*, a series of transgenes containing 0, 5, 30, 83, 125 and 213 CTG repeats, respectively, within the 3'-UTR of *GFP* gene driven by *myo3* promoter were constructed and injected into *lin15*(-/-) strain. Transgenic animals were accordingly designated as *myo3::gfp*, *myo3::gfp*$(CTG)_5$, *myo3::gfp*$(CTG)_{30}$, *myo3::gfp*$(CTG)_{83}$, *myo3::gfp*$(CTG)_{125}$, and *myo3::gfp*$(CTG)_{213}$. Animals expressing CAG repeats of comparable sizes, including *myo3::gfp*$(CAG)_5$, *myo3::gfp*$(CAG)_{30}$, *myo3::gfp*$(CAG)_{83}$, *myo3::gfp*$(CAG)_{120}$ and *myo3::gfp*$(CAG)_{200}$, were also generated in parallel to investigate if the effects are sequence-specific. Except for *myo3::gfp*$(CTG)_{213}$ and *myo3::gfp*$(CAG)_{200}$ animals, the rest of transgenics can produce offspring. The majority of *myo3::gfp*$(CTG)_{213}$ and *myo3::gfp*$(CAG)_{200}$ animals died within two days after injection and did not lay eggs. All laid eggs either stopped developing at three- to four-fold stage or showed retarded growth at L2 stage with abnormal body shape. Since transgene is driven by body wall muscle specific promoter, the uneven distribution of

GFP protein revealed by fluorescent images indicates that the muscle formation is disrupted in the defective embryos and larva. The data demonstrate that transcripts containing $(CUG)_{213}$ and $(CAG)_{200}$ repeats are deleterious to muscle development of *C. elegans*.

The length-dependent toxicity of untranslated CUG and CAG repeats in *C. elegans*

Except for *myo3::gfp(CTG)$_{213}$* and *myo3::gfp(CAG)$_{200}$* transgenics, the animals expressing shorter repeats developed normally. However, the *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals can survive only 5 to 6 days with a relatively small brood size. Besides, they showed deviated irregular locomotory pattern and lower motility rate, significantly deviated from those observed in *myo3::gfp* animals. Further examination of the musculature by staining animals with the filamentous actin (F-actin) marker phalloidin reveals disrupted, wave-like structure in part of muscles in *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals. More than 10% of the examined muscle cells showed disrupted structure in these animals. On the contrary, only few abnormal cells were observed occasionally in animals expressing shorter repeats. The ultra-structure analysis by transmission electron microscope shows that the regular arrangement is disrupted in *myo3::gfp(CTG)$_{125}$* and the dense body disappears in *myo3::gfp(CAG)$_{120}$* abnormal muscles. These findings demonstrate that expansions of both untranslated CTG and CAG repeats are pathogenic *in vivo* with a similar size requirement.

Classical DM1 patients often present the clinical symptoms around adolescence. To investigate if the pathogenic effect of triplet repeats is growth-regulated, synchronized animals at different growth stages were prepared for musculature analysis. The number of abnormal muscle cells keeps increasing during growth. At L2 stage, the ratio of *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals with more than 2 abnormal muscle cells is less than 5% of all examined animals. This ratio increases dramatically (from 10% to more than 50%) during L3 to L4 transition. Therefore, the effect of untranslated CTG and CAG repeats on muscle structure and function is both size- and growth-dependent.

EPG analysis of *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$ C. elegans*

In addition to the muscle phenotype, we also investigated the electropharyngeogram (EPG) of *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals and compared to that of *myo3::gfp* and *myo3::gfp(CTG)$_{5}$* animals. The typical EPG consists of a contraction followed by a relaxation. The contraction potential, which is the large positive transient in the EPG, corresponds to the depolarization of pharyngeal muscle and the onset of contraction. The relaxation potential, which is the large negative transient, corresponds to the repolarization of pharyngeal muscle and the end of action potential. The amplitudes of positive and negative transient were reduced in most examined *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals. This reduction is statistically significant in at least one examined lines, indicating that untranslated CAG repeats may also affect electrophysiological behavior as CTG repeats does.

The effect of CTG and CAG triplet repeats on gene expression

To investigate if the triplet lengths affect gene expression, the GFP fluorescence of transgenic animals was first checked during growth. The fluorescence signal of most *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals turned weak at L3 or L4 stage. In contrast, no significant change was observed for animals expressing shorter repeats. Consistently, Western blot analysis using protein extracts from stage-mixed population shows that the GFP protein expression is greatly decreased by large triplet repeats. To examine if the reduced GFP protein level is due to lower RNA expression, we performed quantitative RT-PCR using total RNA extracted from young adult animals. All examined animals expressed similar levels of

GFP RNA, indicating that large triplet repeats do not interfere with transcription. These results indicate that the effect of large triplet repeats on gene expression is through a posttranscriptional mechanism.

The effect of RNAi-mediated inhibition of *GFP* expression on *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* *C. elegans*

To further confirm the cellular toxicity of transgenes containing large triplet repeats, RNAi-mediated inhibition of *GFP* expression in transgenic animals were performed. To do so, L4440 vectors containing GFP or 125 (CTG/CAG) triplet repeats designated as L4440::GFP and L4440::CTG/CAG125, respectively, were constructed and transformed into HT115 strain. Transgenic animals were then fed with HT115 strains which express GFP or (CUG/CAG) dsRNA upon IPTG induction. The F2 of *myo3::gfp*, *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals were evaluated for RNAi effects. In parallel, animals fed with OP50 strain or HT115 strain expressing no dsRNA (L4440 vector only) were analyzed as a comparison. Western blotting using GFP antibody showed a 40 to 90% inhibition of transgene expression after RNAi treatment. Although the inhibition is incomplete, the abnormal phenotypes of *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* transgenic animals, including life span, brood size, motility, and frequency of animals with abnormal muscle cells, were significant reversed (P<0.01). In addition, the phenotypes of *myo3::gfp* animals were not affected by RNAi treatment, indicating that (CUG/CAG) dsRNA will not cause deleterious effect on *C. elegans*. The data further confirm that the abnormality observed in *myo3::gfp(CTG)$_{125}$* and *myo3::gfp(CAG)$_{120}$* animals is caused by the expression of untranslated triplet repeats.

四、成果自評

我們已建立基因轉殖線蟲之動物模式並進而以 RNAi 方法進一步證實擴增之 CAG 與 CTG 三聯核酸重複序列確會對個體造成傷害。目前我們正探討其他三聯核酸重複序列之影響。

五、參考文獻

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, Sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Johns C, Juvonen V, Johnson K, Harper PS, Shaw DJ, Housman DE (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 68,799-808.

Brais B, Bouchard JP, Xie YG, Rochefort DL, Chretien N, Tome FM, Lafreniere RG, Rommens JM, Uyama E, Nohira O, Blumen S, Korczyn AD, Heutink P, Mathieu J, Duranceau A, Codere F, Fardeau M, Rouleau GA (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet.* 18(2), 164-167.

Buxton J, Schelbourne P, Davies J, Jones C, Van Tongeren T, Aslanidis C, De Jong P, Jansen G, Anvert M et al.. (1992) Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 355, 547-548.

Fu YH, Pizzuti A, Fenwick RG Jr, King J, Rajnarayan S, Dunne PW, Dubel J, Nasser GA, Ashizawa T, de Jong P, Wieringa B, Korneluk R, Perryman MB, Epstein HF, Caskey CT (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255,1256-1258.

Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, Buckler AJ, Harper PS, Housman DE, and Shaw DJ (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy.

Nature 355, 545-546.

Harley HG, Rundle SA, MacMillan JC, Myring J, Brook JD, Crow S, Reardon W, Fenton I, Shaw DJ, and Harper PS (1993) Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am. J. Hum. Genet.* 52, 1164-1174.

La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352(6330), 77-79.

Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 293(5531), 864-867.

Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barcelo J, O'Hoy K, Leblond S, Earle-Macdonald J, de Jong PJ, Wieringa B, Korneluk RG (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255,1253-1255.

Mankodi A, Logigian E, Callahan L, McClain C, White R, Henderson D, Krym M, and Thornton CA (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* 289, 1769-1772.

Monckton DG, Wong LJC, Ashizawa T, and Caskey CT (1995) Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.* 4, 1-8.

Ranum LPW, and Day JW (2004) Myotonic dystrophy: RNA pathogenesis comes into focus. *Am. J. Hum. Genet.* 74, 793-804.

Seznec H, Agbulut O, Sergeant N, Savouret C, Ghestem A, Tabti N, Willer JC, Ourth L, Duros C, Brisson E, Fouquet C, Butler-Browne G, Delacourte A, Junien C, and Gourdon G. (2001) Mice transgenic for the human myotonic dystrophy with expanded CTG repeats display muscular and brain abnormalities. *Hum. Mol. Genet.* 10, 2717-2726.

The international Myotonic dystrophy Consortium (IDMC). (2000) New nomenclature and DNA testing guidelines for myotonic dystrophy type 1 (DM1). *Neurology* 54,1218-1221.

Thornton CA, Johnson K, and Moxley RT (1994) Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann. Neurol.* 35, 104-107.