



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I396742B1

(45)公告日：中華民國 102 (2013) 年 05 月 21 日

(21)申請案號：098136948

(22)申請日：中華民國 98 (2009) 年 10 月 30 日

(51)Int. Cl. : C12N15/89 (2006.01)

C12N15/63 (2006.01)

(71)申請人：中央研究院(中華民國) ACADEMIA SINICA (TW)

臺北市南港區研究院路 2 段 128 號

(72)發明人：陳志毅 CHEN, JYH YIH (TW) ; 彭冠傑 PENG, KUAN CHIEH (TW)

(74)代理人：盧偉銘

(56)參考文獻：

WO 00/49150A1

Clark S, "Zebrafish DNA sequence from clone CH211-129K16 in linkage group 3, complete sequence", NCBI genbank accession no.BX005142, 2003/06/27

Ju B et al., "Recapitulation of fast skeletal muscle development in zebrafish by transgenic expression of GFP under the mylz2 promoter", DEVELOPMENTAL DYNAMICS, VOL.227, P.14-26, 2003

Huang CJ et al., "Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish", DEVELOPMENTAL DYNAMICS, VOL.228, P.30-40, 2003

Zeng ZQ et al., "Faithful expression of living color reporter genes in transgenic medaka under two tissue-specific zebrafish promoters", DEVELOPMENTAL DYNAMICS, VOL.234, P.387-392, 2005

Pan X et al., "Ornamental expression of red fluorescent protein in transgenic founders of white skirt tetra (*Gymnocorymbus ternetzi*)", MARINE BIOTECHNOLOGY, Vol.10, P.497-501, 2008/05/01

審查人員：施雅儀

申請專利範圍項數：10 項 圖式數：7 共 0 頁

(54)名稱

新穎啟動子片段及其應用

A NOVEL PROMOTER AND APPLICATIONS THEREOF

(57)摘要

本發明係關於一種來自斑馬魚肌凝蛋白輕鏈二(mylz 2)基因之新穎啟動子及其應用。本發明之啟動子具提升的啟動子活性，可用於製備轉殖基因動物，特別是螢光魚。

The present invention relates to a novel promoter found in the upstream region of the myosin light chain-2 (mylz2) gene, and applications thereof. The promoter according to the invention exhibits an enhanced promoter activity, which is useful for producing transgenic animals, particularly fluorescent fish.

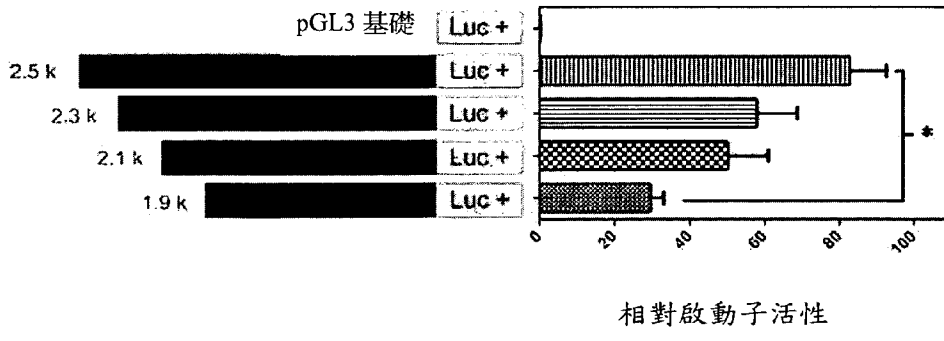


圖 2

公告本

發明專利說明書

(本說明書格式、順序，請勿任意更動，※記號部分請勿填寫)

※申請案號：

9819694

C12N 15/89 (2006.01)

※申請日：

98.10.20

※IPC 分類：

C12N 15/63 (2006.01)

一、發明名稱：(中文/英文)

新穎啟動子片段及其應用

A NOVEL PROMOTER AND APPLICATIONS THEREOF

二、中文發明摘要：

本發明係關於一種來自斑馬魚肌凝蛋白輕鏈二 (*mylz2*) 基因之新穎啟動子及其應用。本發明之啟動子具提升的啟動子活性，可用於製備轉殖基因動物，特別是螢光魚。

三、英文發明摘要：

The present invention relates to a novel promoter found in the upstream region of the myosin light chain-2 (*mylz2*) gene, and applications thereof. The promoter according to the invention exhibits an enhanced promoter activity, which is useful for producing transgenic animals, particularly fluorescent fish.

四、指定代表圖：

(一) 本案指定代表圖為：第 ( 2 ) 圖。

(二) 本代表圖之元件符號簡單說明：

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

(無)

## 六、發明說明：

### 【發明所屬之技術領域】

本發明係關於來自斑馬魚肌凝蛋白輕鏈 2 (mylz2) 基因之新穎啟動子及其應用。

### 【先前技術】

基因轉殖是現今生物科技領域廣泛應用的技術，其係經由遺傳工程技術，透過人為操控方式，將可表現特定基因之 DNA 分子導入動物體細胞或生殖細胞中，以修飾動物之基因表現或賦予新的表現性狀。基因轉殖技術是目前畜養業與水產養殖積極運用的技術之一，特別是針對牛、羊、豬、魚、蝦等經濟動物，可用以改良品種、增加生長率、改良肉質、提高抗病能力及適應力、改變顏色、或提升觀賞價值等。

一般而言，欲導入動物體內之 DNA 分子包含標的基因及用以調控其表現的啟動子序列。選擇適當的啟動子可使標的基因依所欲方式表現，進而達到在轉殖基因動物中之所欲效果，例如，組織專一性表現、誘導性表現、或提升的蛋白質表現能力。

針對魚類，由於基因體學起步較晚，累積的資訊不足，故早期魚類之基因轉殖調控元件大多來自老鼠或病毒，使得轉殖基因魚之體內終究留有病毒來源之 DNA 片段。近年來，研究人員致力於「全魚」(all-fish) 轉殖基因魚之研究，全魚是指所有欲轉殖的基因及所使用的各項調控元件均源自於魚類，使得轉殖成功的魚類得以改變部份基因的調控，以表現出所預期的特性，但不帶有源自其它物種的基因片段。

因此，在此領域中，一直有需要尋求有潛力的啟動子，俾應用於動物之轉殖基因技術，特別是魚類。

### 【發明內容】

因此，在一方面，本發明提供一種分離的核酸分子，其係由 SEQ ID NO: 1 之核苷酸序列所組成。本發明之分離的核酸分子具有啟動子活性。

在另一方面，本發明提供一種重組載體，其包含第一核酸片段，該第一核酸片段係由 SEQ ID NO: 1 之核苷酸序列所組成。典型地，本發明之重組載體尚包括第二核酸片段，其係編碼一蛋白質或多肽，且操作地連接至第一核酸片段。根據本發明，該第二核酸片段係編碼一外源性之蛋白質或多肽。在一具體實例中，本發明之重組載體之第二核酸片段係編碼螢光或冷光蛋白質。

在又一方面，本發明提供一種製備轉殖基因動物之方法，其包含將前述重組載體導入動物細胞中。根據本發明之具體實例，本發明之方法可用於製備轉殖基因魚，特別是螢光魚。

本發明之各個具體實例的細節說明如後。本發明之其他特徵將會經由以下各個具體實例中的詳細說明及申請專利範圍而清楚呈現。

無須進一步的闡述，咸相信本發明所屬技術領域中具有通常知識者基於前述說明即可利用本發明至最廣的程度。因此，可以理解以下的說明僅僅是作為例示說明，而非以任何方式限制本發明範圍。

### 【實施方式】

本發明主要從斑馬魚 mylz2 基因之上游序列分離出一段具提升活性的啟動子之核酸片段，可調控蛋白質或多肽之表現。本發明可利用該核酸片段製備轉殖基因動物，包括但不限於魚類，特別是螢光魚。

除非另有指明，所有在此處使用的技術性和科學性術語具有如同本發明所屬技藝中之通常技術者一般所瞭解的意義。

本文所使用的「一」乙詞，如未特別指明，係指至少一個（一個或一個以上）之數量。

本文所使用的「核酸分子」或「核酸」乙詞是指核苷酸單體之聚合體，其可以是單股或雙股、直線型或環狀、去氧核糖核酸（DNA）或核糖核酸（RNA）。一核苷酸單體包括磷酸基部分、糖基部分及鹼基部分。常見的核苷酸之鹼基部分包括鳥糞嘌呤（G）、腺嘌呤（A）、胞嘧啶（C）、胸腺嘧啶（T）及尿嘧啶（U），其中腺嘌呤與胸腺嘧啶或尿嘧啶配對，以及鳥糞嘌呤與胞嘧啶配對。

本文所使用之「分離的核酸分子」係指從天然生物體來源純化而來的核酸，或以化學合成法或 PCR 擴增技術所製得的核酸。核酸之純化方法、化學合成法及 PCR 擴增技術可參見 *Molecular Cloning: A Laboratory Manual*，第二版，Cold Spring Harbor Laboratory Press，Sambrook J 等人編著，1989，以及 *Current Protocols in Molecular Biology*，Frederick M.A. 等人編著，2001，John Wiley & Sons, Inc.。此一領域具通常知識者可依據習知的技術獲得本發明之核酸分子，詳細方法描述於以下實例中。

本文所使用之「可操作性地連接」或類似用語是指調節性序列與另一編碼特定基因產物之核苷酸序列的連接方式，使得該核苷酸序列接受調節性序列之控制或調節，而在適當條件下表現出該基因產物。

本文所使用的「重組」係用以描述核酸分子含有非天然連接的序列。重組核酸分子可為載體形式，其可包括一或多個編碼基因之結構序列及用於表現該基因之調節性序列，例如，啟動子序列。載體之種類包括但不限於質體、黏質體、噬菌體、YAC 或 PAC 等。重組載體可用於表現或複製所導入的基因序列。典型地，在重組載體中，編碼基因之結構序列係操作性地連結於調節性序列，使得重組載體導入適當宿主中時，編碼基因之結構序列在該調節性序列之調控下表

現。調節性序列可包括但不限於啟動子序列、起始密碼子、複製起始點或分泌訊息序列等。

本文所使用的「核酸分子編碼．．．」或其類似用語是指，該核酸分子帶有一或多個編碼特定基因產物之核苷酸序列，而可在適當條件下作為模板表現出該基因產物。應瞭解的是，由於基因密碼（由三個核苷酸所組成）之簡併性，故不同的基因密碼可編碼相同的胺基酸殘基，例如，GAU 與 GAC，兩者在核苷酸序列上不同，但均編碼相同胺基酸殘基，即天門冬酸 (Asp)。針對相同的多肽產物，所屬領域具通常知識者可運用例行技術及已知的基因密碼偏好性 (codon usage)，獲得對應的核苷酸序列及其簡併序列。因此，本文所述編碼特定基因產物之核苷酸序列尚包括其簡併序列（可編碼相同基因產物）。

在一方面，本發明提供一種分離的核酸分子，其係由 SEQ ID NO: 1 之核苷酸序列所組成。具體而言，本發明之核酸分子係分離自斑馬魚 mylz2 基因之上游序列，其長度為 2.5kb 並具有提升的啟動子活性，可用於表現操作性地連接於其上的基因序列。根據本發明實施例，該核酸分子與其他啟動子片段相比，具較佳之啟動子活性；尤其較 1.9 kb 啟動子片段之活性高出約 2.8 倍之多(如實例 3 所列)。

在另一方面，本發明提供一種重組載體，其包含第一核酸片段，該第一核酸片段係由 SEQ ID NO: 1 之核苷酸序列所組成。典型地，本發明之重組載體進一步包含第二核酸片段，其編碼蛋白質且操作性地連結至第一核酸片段。根據本發明，該第二核酸片段表現係由第一核酸片段調控。根據本發明，該第二核酸片段可為修飾動物之基因表現或賦予新的表現性狀，例如，增加生長率、改良肉質、提高抗病能力及適應力、改變顏色、或提升觀賞價值等。本發明一具體實施態樣中，第二核酸片段係編碼一外源性蛋白質或多肽，例如該外源蛋白質可為冷光或螢光蛋白質俾製成觀賞魚，或該外源



蛋白質可為抗病源蛋白等俾製成抗病魚增加魚類生存活力。

根據本發明，冷光蛋白包括但不限於螢火蟲冷光酶及水母冷光酶；以及螢光蛋白包括但不限於綠螢光蛋白 (GFP)、藍綠螢光蛋白 (CFP)、紅螢光蛋白 (RFP)、黃螢光蛋白 (YFP) 及其修飾形式，例如，強化型綠螢光蛋白 (EGFP)、強化型藍綠螢光蛋白 (ECFP)、強化型紅螢光蛋白 (ERFP) 及強化型黃螢光蛋白 (EYFP)。此等蛋白之胺基酸及核苷酸序列已為此領域所習知者。例如，紅螢光蛋白具有 SEQ ID NO: 2 之核苷酸序列及 SEQ ID NO: 3 之胺基酸序列 (圖 4)。

在一特定實例中，本發明之重組載體包含 SEQ ID NO: 4 之核苷酸序列，其涵蓋 SEQ ID NO: 1 之核苷酸序列 (啟動子) 及編碼紅螢光蛋白之 SEQ ID NO: 2 核苷酸序列。

更進一步說明，本發明之載體尚可視需要地包括其他具有特定功能之序列，例如，複製起始點、終止序列、供篩選用途之標記基因，例如，抗生素抗性基因等。此一領域具通常知識者可依現有知識選擇適當的序列，使用習知的遺傳工程技術構築本發明之載體。構築重組載體所需使用的遺傳工程技術，例如，聚合酶連鎖反應 (PCR) 擴增技術、核酸純化、載體及核酸片段之酵素處理及定序技術，可參見前述 *Molecular Cloning: A Laboratory Manual* 及 *Current Protocols in Molecular Biology* (同上)。具體的載體構築細節說明於以下的實例中。

因此，在又一方面，本發明提供一種製備轉殖基因動物之方法，其包含將此處所述之重組載體導入動物細胞中。

根據本發明，此處所述「轉殖基因動物」包括但不限於牛、羊、豬、魚、蝦等不包含人類之動物。特定而言，根據本發明之轉殖基因動物是魚類，更特定而言，為科學上或商業上有應用價值的魚種，包括但不限於斑馬魚 (Zebrafish, 如 *Danio rerio*)、神仙魚 (angelfish, 如七彩神仙 (discuss)、*Pterophyllum*)、青鱗魚 (medaka)、金魚 (goldfish)、鯉魚

(carp)、錦鯉 (koi)、泥鰍 (loach)、吳郭魚 (tilapia)、玻璃魚 (glassfish)、鯰魚 (catfish)、鰻 (eel)、紅蓮燈 (tetra)、蝦虎 (goby)、攀鱸科淡水魚 (gourami)、孔雀魚 (guppy)、劍魚 (Xiphophorus)、茉莉魚 (Molly fish) 及鯰魚 (pangasius)。

在一特定實例中，根據本發明之轉殖基因動物是斑馬魚。

本發明之方法可利用已知的轉殖基因標準技術，例如顯微注射 (microinjection)、電破法 (electroporation)、病毒載體法 (virus infection) 及精子載體法 (sperm-mediated gene transfer) 等方法，迫使本發明之核酸分子進入動物之配子、受精卵、胚胎或體細胞的細胞核或細胞質內，使本發明核酸分子能在胚胎或細胞中繼續複製，進而在個體及其後代細胞上表現，製得轉殖基因動物。

在一特定實例中，根據本發明之轉殖基因動物是斑馬魚，其係使用顯微注射技術將重組載體導入魚卵細胞中而製得。進一步說明，使用毛細針吸取適當體積的重組載體，將其注射於單細胞時期之斑馬魚受精卵，將注射完之魚卵集中於乾淨培養皿裡，至於 28°C 培養箱培養，約 48 至 72 小時孵出魚苗。

此一領域已發展出各種習知方法辨別或篩選帶有轉殖基因的動物。例如，設計可辨認轉殖基因之探針，以北方墨漬法 (Northern blotting) 或南方墨漬法 (Southern blotting) 來偵測候選動物中是否具有轉殖基因的序列，或以 PCR 針對特定核酸序列放大的方式進行偵測。特定而言，表現螢光蛋白之動物可直接用眼睛觀察。

在一特定實例中，本發明之方法係將編碼螢光蛋白之重組載體導入受精魚卵中，經培養後孵化，直接用眼睛篩選發出螢光的轉殖基因魚。

本發明之各個具體實例的細節說明如後。本發明之技術特徵將會經由以下各個具體實例中的詳細說明及申請專利範圍而更清楚呈現。

### 實例 1：基因組 DNA 之萃取

本實驗採用斑馬魚 (Zebrafish, *Danio rerio*) 為材料，使用商業套組 PureLink™ Genomic DNA Mini Kit (Invitrogen, 美國) 萃取基因組 DNA。簡言之，將組織放入微量試管之中，加入 PureLink™ 基因組消化緩衝液和蛋白酶 K，於 55°C 條件下分解組織，其間偶爾的震盪，均勻試劑，直至組織完全分解，過程約 1 至 4 小時。之後在室溫下離心，取上清液置於新的微量試管，加入 RNase A，輕微的震盪，靜置室溫 2 分鐘。加入 200 µl 的 PureLink™ 基因組瓦解/結合緩衝液，混合均勻後，再加入酒精 (96-100%)，震盪混合。將溶解產物以離心的方式通過管柱過濾，清洗後，加入二次水或 PureLink™ 基因組沖提緩衝液，靜置後離心，獲得純化的基因組 DNA。

### 實例 2：啟動子片段之選殖

設計正反引子，交由生工有限公司及台灣生醫聯盟合成，引子序列如下：

引子名稱	序列
5' mylz2 2.5k <i>Xho</i> I	CGCTCGAGATGCTGTGAAGTATTCTCTACTT (SEQ ID NO: 5)
3' mylz2 2.5k <i>Nhe</i> I	CGGCTAGCGTAGTGTCCTGTACTTGAGGGGC (SEQ ID NO: 6)

以斑馬魚基因組 DNA 為模板，在 PCR 離心管中添加上述正反引子、*Taq* DNA 聚合酶、緩衝液、dNTPs 等試劑，含量如下：

試劑	含量
基因組 DNA	100 ng
10X ThermoPol 反應緩衝液	2.5 µl
10 mM dNTP	2 µl
50 µM 引子(5' mylz2 2.5k <i>Xho</i> I)	0.25 µl

50 $\mu$ M 引子(3' <i>mylz2</i> 2.5k <i>NheI</i> )	0.25 $\mu$ l
<i>Taq</i> DNA 聚合酶 (5 U/ $\mu$ L)	0.25 $\mu$ l
ddH <sub>2</sub> O	加滿至總體積為 25 $\mu$ l

將上述 PCR 離心管移至 PCR 反應器 (Applied Biosystems 2700, 美國), 進行 PCR 反應, 其條件為: 起始變性反應為 94°C 5 分鐘, 變性、黏合及延伸反應之循環為 94°C 30 秒、55°C 30 秒及 72°C 2 分鐘, 重複 35 個循環, 最後延伸反應為 72°C 10 分鐘, 最後置於 4°C 終止反應。PCR 反應結束後, 以瓊脂膠體電泳分析 PCR 產物, 再將含有預測大小的 PCR 產物片段之膠體切下, 以商業套組 Gel/PCR DNA Fragments Extraction Kit (Geneaid, Taiwan), 獲得純化的 PCR 產物。

將純化的 PCR 產物選殖至 pCRII-TOPO-TA 載體 (Invitrogen), 先經由轉殖菌株之抗生素篩選, 再由台北基龍米克斯生物科技股份有限公司以核酸序列自動定序儀 (autosequencer, ABI) 進行定序工作, 從 5' 端及 3' 端進行至少兩次定序分析, 選出品系, 確認其含有重組載體 pCRII-*mylz2*-2.5k, 該重組載體帶有核苷酸序列為 SEQ ID NO: 1 之啟動子片段, 下稱「2.5kb 啟動子片段」。圖 1 顯示 2.5kb 啟動子片段之核苷酸序列 (-1 至 -2504)。

### 實例 3: 冷光酶重組載體之構築及啟動子活性分析

以上述重組載體 pCRII-*mylz2* 2.5k 為模板, 分別使用下列 5' 端引子之其中一種, 加上 3' 端引子, 以 PCR 增幅出 2.5 kb 啟動子片段及其截短片段 (2.3 kb、2.1kb 及 1.9kb):

引子名稱	序列
<i>mylz2</i> 2.5K 5' <i>SacI</i>	CGGAGCTCATGCTGTGAAGTATTCTCTA (SEQ ID NO: 7)
<i>mylz2</i> 2.3K 5' <i>SacI</i>	CGGAGCTCTACTAATTGATGATTAGTTT (SEQ ID NO: 8)

mylz2 2.1K 5' <i>SacI</i>	CGGAGCTCTTGCCAGTTT TAGTTCAGTCA (SEQ ID NO: 9)
mylz2 1.9K 5' <i>SacI</i>	CGGAGCTCGTGTATTTTTTATGCAGCGG (SEQ ID NO: 10)
mylz2 promoter 3' <i>XhoI</i>	CGCTCGAGGTAGTGTCTGCTGTA CTTGAGG (SEQ ID NO: 11)

將各種長度之啟動子片段分別選殖至 pGL3 基礎冷光報導載體 (Promega, 美國), 獲得含有不同長度之啟動子片段之重組 pGL3 載體。使用表現水母冷光酶基因之 pRL-TK 載體做為內部對照組, 以 1:1 之比例, 分別與前述重組 pGL3 載體 (含有不同長度之啟動子片段) 或 pGL3 基礎載體 (不含啟動子片段) 混合後, 以顯微注射法, 導入單細胞時期之斑馬魚受精卵。使受精卵在 28°C 培育 72 小時, 以商業套組 Dual-Luciferase® Reporter Assay System (Promega, 美國) 測量啟動子活性。

簡言之, 每組取 10 個健康胚胎, 加入 50 µl 被動裂解緩衝液 (passive lysis buffer), 在冰上磨碎。離心 10 分鐘, 取上清液與 50 µl 冷光酶分析試劑 II (LAR II) 混合, 使用冷光分析儀 (Fluoroskan Ascent FL luminometer, Thermo Labsystems, 美國) 測量發光。圖 2 顯示啟動子活性分析結果。數據係以平均數值±SEM 呈現, 共進行三重複實驗, 使用 SAS 統計程式 (SAS Institute, 美國), 統計上的顯著性設定於  $p < 0.05$  或  $p < 0.01$ , 所有數據係以單因子變異數分析 (ANOVA) 進行分析而得。

如圖 2 所示, 相較於不具啟動子片段的對照組載體 (pGL3 基礎載體) 以及含有其他長度之啟動子片段的重組 pGL3 載體 (2.3 kb、2.1kb 及 1.9kb), 本發明之含有 2.5kb 啟動子片段之重組 pGL3 載體展現顯著提升的啟動子活性, 其中比較 2.5 kb 啟動子片段及 1.9 kb 啟動子片段, 2.5 kb 啟動子片段之活性約高出 2.8 倍之多。

#### 實例 4：轉殖基因魚（螢光魚）之製備

從前述重組載體 pCRII-mylz2-2.5k 切下 2.5kb 啟動子片段，接入可表現紅螢光之 pDsRed2-1 載體（Clontech，美國），挑選正確品系，獲得 pmylz2 2.5K-DsRed 重組載體（圖 3）。

使用毛細針吸取適當體積的重組載體（濃度不超過 30 ng/ $\mu$ l）和追蹤染劑（最終濃度為約 2% 的酚紅），對單細胞時期之斑馬魚受精卵進行顯微注射。將注射完之魚卵集中於乾淨培養皿裡，加入曝氣水於 28°C 培養箱培養。每天檢查發育情形並挑出發育失敗之魚卵，待 48-72 小時孵出魚苗。

養至成魚之後，依外觀表現從中挑選 6 隻轉殖基因候選魚，分別為：雄魚 1 號、雄魚 2 號、雄魚 3 號、雄魚 5 號、雄魚 6 號、和雌魚 1 號，再個別與野生型斑馬魚進行交配。轉殖基因候選魚之螢光魚子代產出率分別為：18.23%、18.06%、0%、0%、0%和 5.88%，其中又以雌魚 1 號之 F1 的螢光表現最強。繼而從雌魚 1 號之 F1 中篩選出達目標之表現的個體，養成後再與野生型斑馬魚交配，產生 F2，隨後以 F2 自交，純化並穩定品系，至少可繼續繁衍至第五代，命名此穩定品系為 pm2.5k-DsRed Tg1。該品系在腹部有明顯螢光表現，更甚者在吻部亦有明顯螢光表現；相較之下，含有 1.9 kb 啟動子片段之轉殖基因斑馬魚，腹部無明顯螢光表現，呈白色狀（Gong et al., 2003）。

綜上，本發明之 2.5 kb 啟動子片段源自斑馬魚 mylz2 基因，相較於其他長度之啟動子片段，具顯著提升的啟動子活性。本發明之啟動子片段可用於轉殖基因動物之製備，特別對於水產動物之基因轉殖研究有極大助益，更特別是在螢光魚之製備上有顯著提升螢光表現之效果，對於觀賞魚之產業有相當貢獻。

無須進一步的闡述，咸相信本發明所屬技術領域中具有通常知識者基於前述說明即可利用本發明至最廣的程度。因此，

可以理解以下的說明僅僅是作為例示說明之用，而非以任何方式限制其餘的揭露內容。

### 【圖式簡單說明】

圖 1A 及 1B 顯示 2.5kb 啟動子片段之核苷酸序列 (SEQ ID NO: 1)。數字代表轉譯起始位置 ATG 向上游計算之核苷酸位置。

圖 2 顯示本發明之 2.5kb 啟動子片段之冷光酶活性分析結果。

圖 3 顯示根據本發明之用於製備轉殖基因魚之 pm2.5K-DsRed 重組載體之結構圖。

圖 4 顯示 pm2.5K-DsRed 重組載體編碼之紅螢光蛋白之核苷酸序列 (SEQ ID NO: 2) 及胺基酸序列 (SEQ ID NO: 3)。

圖 5A 及 5B 顯示 pm2.5K-DsRed 重組載體編碼之啟動子至紅螢光蛋白之核苷酸序列 (SEQ ID NO: 4)。

## 序列表

- <110> 中央研究院  
 <120> 新穎啟動子片段及其應用  
 <160> 11  
 <170> PatentIn version 3.3  
 <210> 1  
 <211> 2504  
 <212> DNA  
 <213> 斑馬魚 (*Danio rerio*)  
 <220>  
 <223> 2.5 kb 啟動子  
 <400> 1

atgctgtgaa	gtattctcta	cttatcttaa	ttctctaaaa	ataatgcaaa	agtattcctt	60
aaaatgttaa	atgaagggat	atcttcaaag	ccaatcgact	tttagaatag	ataatgttta	120
attgactaac	agcaattcag	ctacaattca	agtaatgttc	tgaagataga	tatgagttaa	180
caaatctctg	gacaactaca	tactaattga	tgattagttt	aaacctatta	ttttacatta	240
cagggagtat	actacaatat	aaatacgata	aaacttgtga	gcatgtatta	aagcaattaa	300
taattcctca	atcaaatac	accccatatg	tcccatgctc	agcatatatg	agtacacccc	360
tctagcctga	tctcaagaga	aaacgcgagt	atcttatggt	ttgccagttt	agttcagtca	420
ttgggggatg	agcaaatcat	actaaaatgt	acgaattagt	togtacgaat	tagccactaa	480
atgaaaaagt	tacaaatttg	tgttgacccc	tcacaaatct	gtcttcattc	atcttctttt	540
aggcttagtc	cctttattaa	tcaggattcg	ccacagagga	atgagccacc	aactcatcca	600
gtgtattttt	tatgcagcgg	atgccatcc	agttgcagcc	ctacactggg	aaacacccaa	660
atctgtcttt	taaattaata	tttttaatag	gaagctatac	aatgttatat	ttgtgcatat	720
acattagatt	agtcagtaca	ctctcagaaa	taaaggtgcg	caagctgtca	ctgcggtggt	780
accttttcaa	aaggtacata	tttgtactta	aaggggccat	attgatacct	caaaagtata	840
tattagtgcc	taaaaatttt	aagagaaaca	cttttgtact	ttttaggtac	taatataac	900
ccttgagggt	ttattatgga	cctttaaggt	acaaattttt	atcatttgga	aaggtaccac	960
cccagtgaca	gtttacaccc	tttatttctg	agagtgaaagc	taaatctgga	gcttatttaa	1020
caaaataact	tatgataacg	gtccaaaac	tactacaccc	atatttatat	gttattgaaa	1080
aatattaat	acaaatttta	aaacagggaa	aatcaagag	aagcaaaaca	atggaaaaaa	1140
aattgttcaa	aattttagg	ttgtaatttt	tttgttgcaa	tatttagctt	tatttaaatt	1200
gttttatctt	tcaattttgt	gactaaaata	ttattttaat	aaatatagct	gtttaataaa	1260
tctgttttgt	ttaaatgcac	caaaatacat	tgcttatatt	cactgagaga	tgagaaaaaa	1320
tattcatttt	caaaatgggc	tgtactcgat	tactctgagc	actgtaatta	tctgtttttc	1380
taggtactgt	aaaataatgc	taatcttaaa	caattttgga	acaagaagga	gcaaaacaag	1440
ccgacagact	cacaggatat	caaattagtc	taaagaaaga	agggaatgca	caacaaactc	1500
aagggggaca	aaacagtgac	tgatgaatta	gacaagaaaa	agagagcaag	gagcgtcaa	1560
gattgtttag	ctattttggg	cacccacagc	tgctccttat	gcctgccttc	caaaaaaaa	1620
actgtcttaa	gcctcaaatt	tctcttcatg	aggggtccaac	atcaaccact	cagagggctg	1680
tagtgtgctg	accatctaaa	aactgggaaa	aaggggtaat	tacgtgcttg	tccacagggc	1740
agcttcccac	aatggcacc	tcacagtcac	tgaagtgacc	gggtgaggtc	ataggtcgat	1800
cggcagagag	agagagagag	agagagagag	agtgctgaat	atgggcatgg	ccatgtggac	1860
gggtgtggtg	ggggcacttg	aaccgaaatc	ttacagcatc	actatactag	gaaaagcatt	1920
aaaacctatt	ttcgattagg	gctgatttga	aataaggggt	aagacaccag	aacgtcctct	1980
tatatatcag	ccaggggcgt	gaacaaatat	gaacaacata	atcattggct	caaaaaatct	2040
ctggattgaa	atccatcagg	atctatcact	gcaaccctcc	ccatccaaca	agtaatcctg	2100
caagataagc	cagtattgat	ctgctgctaa	tctaacttta	ggggagggag	ttgcgtcccc	2160
tttagactct	gtggctacag	ctcattcatt	tcaaatggag	ttatgtgatt	gtatgaagct	2220
caaacagtcc	cttacgtccc	catgtcctta	ttagtcaacg	cgagacatgc	agggcgtctg	2280



catcagtatc	agattcatcc	cattccaaga	ctccaatagc	tatttctgag	cactgtaaga	2340
tgatagtaca	tcccagccgg	tgtccctcca	tcactttccc	cctacctcat	agttttgcct	2400
ctttctctct	cggctctgcta	tttcccaaac	ctcacttaag	gttgggtcta	taattcgcaa	2460
ggggccttcg	tcagtatata	agcccctcaa	gtacaggaca	ctac		2504

<210> 2  
 <211> 678  
 <212> DNA  
 <213>

<220>  
 <223> 紅螢光核苷酸序列

<400> 2

atggcctcct	ccgagaacgt	catcaccgag	ttcatgcgct	tcaaggtgcg	catggagggc	60
accgtgaacg	gccacgagtt	cgagatcgag	ggcgagggcg	agggccgccc	ctacgagggc	120
cacaacaccg	tgaagctgaa	ggtgaccaag	ggcggccccc	tgcccttcgc	ctgggacatc	180
ctgtccccc	agttccagta	cggctccaag	gtgtacgtga	agcaccgcc	cgacatcccc	240
gactacaaga	agctgtcctt	ccccgagggc	ttcaagtggg	agcgcgtgat	gaacttcgag	300
gacggcggcg	tggcgaccgt	gaccaggac	tctcctctgc	aggacggctg	cttcatctac	360
aaggtgaagt	tcatcggcgt	gaacttcccc	tccgacggcc	ccgtgatgca	gaagaagacc	420
atgggctggg	aggcctccac	cgagcgcctg	tacccccgcg	acggcgtgct	gaagggcgag	480
accacaagg	ccctgaagct	gaaggacggc	ggccactacc	tggtggagtt	caagtccatc	540
tacatggcca	agaagcccgt	gcagctgccc	ggctactact	acgtggacgc	caagctggac	600
atcacctccc	acaacgagga	ctacaccatc	gtggagcagt	acgagcgcac	cgagggccgc	660
caccacctgt	tctctgtag					678

<210> 3  
 <211> 225  
 <212> 蛋白質  
 <213>

<220>  
 <223> 紅螢光胺基酸序列

<400> 3

MASSENVITE	FMRFKVRMEG	TVNGHEFEIE	GEGEGRPYEG	HNTVKLKVTK	GGPLPFAWDI	60
LSPQFQYGSK	VYVKHPADIP	DYKKSFPPEG	FKWERVMNFE	DGGVATVTQD	SSLQDGCFIY	120
KVKFIGVNFP	SDGPVMQKKT	MGWEASTERL	YPRDGVKGE	THKALKLKD	GHYLVEFKSI	180
YMAKKPVQLP	GYYYVDAKLD	ITSHNEDYTI	VEQYERTEGR	HHLFL		225

<210> 4  
 <211> 3261  
 <212> DNA  
 <213>

<220>  
 <223> 人工序列

<400> 4

gaattcgccc	ttcgctcgag	atgctgtgaa	gtattctcta	cttatcttaa	ttctctaaaa	60
ataatgcaaa	agtattcttt	aaaatgttaa	atgaagggat	attttcaaag	ccaatcgact	120
tttagaatag	ataatgttta	attgactaac	agcaattcag	ctacaattca	agtaatgttc	180
tgaagataga	tatgagttaa	caaatctctg	gacaactaca	tactaattga	tgattagttt	240
aaacctatta	ttttacatta	cagggagtat	actacaatat	aaatacgata	aaacttgtga	300
gcatgtatta	aagcaattaa	taattcctca	atcaaatac	accccatatg	tcccatgctc	360
agcatatatg	agtacacccc	tctagcctga	tctcaagaga	aaacgcgagt	attttatggt	420
ttgccagttt	agttcagtca	ttgggggatg	agcaaatcat	actaaaatgt	acgaattagt	480
tcgtacgaat	tagccactaa	atgaaaaagt	tacaaatttg	tgttgacccc	tcacaaatct	540
gtcttcattc	attttctttt	aggcttagtc	cctttattaa	tcaggattcg	ccacagagga	600
atgagccacc	aactcatcca	gtgtattttt	tatgcagcgg	atgcccatcc	agttgcagcc	660
ctacactggg	aaacacccaa	atctgtcttt	taaattaata	tttttaatag	gaagctatac	720
aatgttatat	ttgtgcatat	acattagatt	agtcagtaca	ctctcagaaa	taaaggtgcg	780
caagctgtca	ctgcggtggt	accttttcaa	aaggtacata	tttgtaacta	aaggggccat	840
attgatacct	caaaggtata	tactagtgcc	taaaaatttt	aagagaaaca	cttttgtact	900
ttttaggtac	taatataac	ccttgagggt	ttattatgga	cctttaaggt	acaaattttt	960
atcagttgga	aaggtaccac	cccagtgaca	gtttacaccc	tttatttctg	agagtgaagc	1020
taaatctgga	gcttatttaa	caaaataact	tatgataacg	gtccaaaaac	tactacaccc	1080
atatttatat	gttattgaaa	aatattaat	acaaatttta	aaacagggaa	aatcaagag	1140
aagcaaaaca	atggaaaaaa	aattgttcaa	aatttgtagg	ttgtaatttt	tttgttgcaa	1200
tatttagctt	tatttaaatt	gttttatctt	tcaattttgt	gactaaaata	ttattttaat	1260
aaatatagct	gtttaataaa	tctgttttgt	ttaaatgcac	caaaatacat	tgcttatatt	1320
cactgagaga	tggggaaaaa	tattcatttt	caaaatgggc	tgtactcgat	tactctgagc	1380
actgtaatta	tctgtttttc	taggtactgt	aaaataatgc	taatcttaa	caattttgga	1440
acaagaagga	gcaaaacaag	ccgacagact	cacaggatat	caaattagtc	taaagaaaga	1500
agggaatgca	caacaaactc	aagggggaca	aaacagtgac	tgatgaatta	gacaagaaaa	1560
agagagcaag	gagcgtcaa	gattgttttag	ctattttggg	cacccacagc	tgttccttat	1620
gctgccttc	ccaaaaaaa	ctgtcttaag	cctcaaattt	ctcttcatga	gggtccaaca	1680
tcaaccactc	agagggctgt	agtgtgctga	ccatctaaaa	actgggaaaa	aggggtaatt	1740
acgtgcttgt	ccacagggca	gcttcccaca	aatggcacct	cacagtcact	gaagtgaccg	1800
ggtgaggtea	taggtcgatc	ggcagagaga	gagagagaga	gagagagtgc	tgaatatggg	1860
catggccatg	tggacgggtg	tgggtggggc	acttgaaccg	aatcttaca	gcatcactat	1920
actagaaaa	gcattaaaac	ctattttcga	ttagggctga	tttgaaataa	gggttaagac	1980
accagaacgt	cctcttatat	atcagccagg	ggcgtgaata	aatatgaaca	acataatcat	2040
tggctcaaaa	aatctctgga	ttgaaatcca	tcaggatcta	tcactgcaac	cctccccatc	2100
caacaagtaa	tcttgcaaga	taagccagta	ttgatctgct	gctaatactaa	ctttagggga	2160
gggagttgcg	tcccctttag	actctgtggc	tacagctcat	tcatttcaaa	ttgagttatg	2220
tgattgtatg	aagctcaaac	agtcccttac	gtcccatatg	ccttattagt	caacgcgaga	2280
catgcaggcc	gctgccatca	gtatcagatt	catcccatc	caagactcca	atagctattt	2340
ctgagcactg	taagatgata	gtacatccca	gcccgtgtcc	ctccatcact	ttccccctac	2400
ctcatagttt	tgctcttttc	tctctcggtc	tgctatttcc	caaacctcac	ttaaggttgg	2460
gtctataatc	agcaaggggc	cttcgtcagt	atataagccc	ctcaagtaca	ggacactacg	2520
ctagccgaag	ggcgaattct	gcagtcgacg	gtaccgcggg	cccgggatcc	accggtcgcc	2580
accatggcct	cctccgagaa	cgtcatcacc	gagttcatgc	gcttcaaggt	gcgcatggag	2640

ggcaccgtga acggccacga gttcgagatc gagggcgagg gcgagggccg cccctacgag 2700  
 ggccacaaca ccgtgaagct gaaggtgacc aagggcggcc ccctgccctt cgcctgggac 2760  
 atcctgtccc ccagttcca gtacggctcc aaggtgtacg tgaagcacc cgccgacatc 2820  
 cccgactaca agaagctgtc cttccccgag ggcttcaagt gggagcgcgt gatgaacttc 2880  
 gaggacggcg gcgtggcgac cgtgaccag gactcctccc tgcaggacgg ctgcttcac 2940  
 tacaaggtga agttcatcgg cgtgaacttc ccctccgacg gccccgtgat gcagaagaag 3000  
 accatgggct gggaggcctc caccgagcgc ctgtaccccc gcgacggcgt gctgaagggc 3060  
 gagaccaca aggcctgaa gctgaaggac ggcggccact acctggtgga gttcaagtcc 3120  
 atctacatgg ccaagaagcc cgtgcagctg cccggctact actacgtgga cgccaagctg 3180  
 gacatcacct ccacaacga ggactacacc atcgtggagc agtacgagcg caccgagggc 3240  
 cgccaccacc tgttctgta g 3261

<210> 5  
 <211> 31  
 <212> DNA  
 <213> 人工合成

<220>  
 <223> DNA 引子 (5' mylz2 2.5k XhoI)

<400> 5  
 cgctcgagat gctgtgaagt attctctact t 31

<210> 6  
 <211> 31  
 <212> DNA  
 <213> 人工合成

<220>  
 <223> DNA 引子 (3' mylz2 2.5k NheI)

<400> 6  
 cggtcagcgt agtgcctctgt acttgagggg c 31

<210> 7  
 <211> 28  
 <212> DNA  
 <213> 人工合成

<220>  
 <223> DNA 引子 (mylz2 2.5K 5' SacI)

<400> 7  
 cggagctcat gctgtgaagt attctcta 28

<210> 8  
<211> 28  
<212> DNA  
<213> 人工合成

<220>  
<223> DNA引子 (mylz2 2.3K 5' SacI)

<400> 8  
cggagctcta ctaattgatg attagttt 28

<210> 9  
<211> 28  
<212> DNA  
<213> 人工合成

<220>  
<223> DNA引子 (mylz2 2.1K 5' SacI)

<400> 9  
cggagctctt gccagtttag ttcagtca 28

<210> 10  
<211> 28  
<212> DNA  
<213> 人工合成

<220>  
<223> DNA引子 (mylz2 1.9K 5' SacI)

<400> 10  
cggagctcgt gtatttttta tgcagcgg 28

<210> 11  
<211> 28  
<212> DNA  
<213> 人工合成

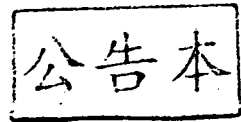
<220>  
<223> DNA引子 (mylz2 promoter 3' XhoI)

<400> 11  
cgctcgaggt agtgcctctgt acttgagg 28

## 七、申請專利範圍：

公告本

1. 一種分離的核酸分子，其係由 SEQ ID NO: 1 之核苷酸序列所組成，其具有啟動子活性。
2. 一種重組載體，其包含第一核酸片段，該第一核酸片段係由 SEQ ID NO: 1 之核苷酸序列所組成。
3. 根據申請專利範圍第 2 項之重組載體，其進一步包含第二核酸片段，其編碼蛋白質且操作地連結至第一核酸片段。
4. 根據申請專利範圍第 3 項之重組載體，其中第二核酸片段係編碼外源蛋白質或多肽。
5. 根據申請專利範圍第 3 項之重組載體，其中第二核酸片段係編碼冷光或螢光蛋白。
6. 根據申請專利範圍第 3 項之重組載體，其中第二核酸片段係由 SEQ ID NO: 2 之核苷酸序列所組成。
7. 根據申請專利範圍第 4 項之重組載體，其包含 SEQ ID NO: 4 之核苷酸序列。
8. 一種製備轉殖基因魚之方法，其包含將申請專利範圍第 3 至 7 項中任一項之重組載體導入魚的配子、受精卵、胚胎或體細胞的細胞核或細胞質內，使該重組載體所含之核酸分子在胚胎或細胞中繼續複製，進而在個體及其後代細胞上表現。
9. 根據申請專利範圍第 8 項之方法，其係使用顯微注射技術將該重組載體導入魚的受精卵中。
10. 根據申請專利範圍第 8 項之方法，其中該魚係斑馬魚。



八、圖式：

```

-2500          -2480          -2460
ATGCTGTGAAGTATTCTCTACTTATCTTAATTCTCTAAAAATAATGCAAAAGTATTCTTT
-2440          -2420          -2400
AAAATGTTAAATGAAGGGATATTTCAAAGCCAATCGACTTTTAGAATAGATAATGTTTA
-2380          -2360          -2340
ATTGACTAACAGCAATTCAGCTACAATTCAGTAATGTTCTGAAGATAGATATGAGTTAA
-2320          -2300          -2280
CAAATCTCTGGACAACCTACATACTAATTGATGATTAGTTTAAACCTATTATTTTACATTA
-2260          -2240          -2220
CAGGGAGTATACTACAATATAAATACGATAAAAACCTTGTGAGCATGTATTAAGCAATTA
-2200          -2180          -2160
TAATTCCTCAATCAAATATCACCCCATATGTCCCATGCTCAGCATATATGAGTACACCC
-2140          -2120          -2100
TCTAGCCTGATCTCAAGAGAAAACGCGAGTATTTTATGTTTTGCCAGTTTAGTTCAGTCA
-2080          -2060          -2040
TTGGGGGATGAGCAAATCATACTAAAATGTACGAATTAGTTCGTACGAATTAGCCACTAA
-2020          -2000          -1980
ATGAAAAAGTTACAAATTTGTGTTGACCCCTCACAAATCTGTCTTCATTCATTTTCTTTT
-1960          -1940          -1920
AGGCTTAGTCCCTTTATTAATCAGGATTCGCCACAGAGGAATGAGCCACCAACTCATCCA
-1900          -1880          -1860
GTGTATTTTTTATGCAGCGGATGCCCATCCAGTTGCAGCCCTACACTGGGAAACACCCAA
-1840          -1820          -1800
ATCTGTCTTTTAAATTAATATTTTTAATAGGAAGCTATAACAATGTTATATTTGTGCATAT
-1780          -1760          -1740
ACATTAGATTAGTCAGTACACTCTCAGAAATAAAGGTGCGCAAGCTGTCACTGCGGTGGT
-1720          -1700          -1680
ACCTTTTCAAAGGTACATATTTGTACTTAAAGGGTCCATATTGATACCTCAAAGTATA
-1660          -1640          -1620
TATTAGTGCCATAAAATTTAAGAGAAACACTTTTGTACTTTTTAGGTAATAATATATAC
-1600          -1580          -1560
CCTTGAGGTTTTATTATGGACCTTTAAGGTACAAATTTTATCATTTGGAAAGGTACCAC
-1540          -1520          -1500
CCCAGTGACAGTTTACACCCTTTATTTCTGAGAGTGAAGCTAAATCTGGAGCTTATTTAA
-1480          -1460          -1440
CAAATAAAGTATGATAACGGTCCAAAAACTACTACACCCATATTTATATGTTATTGAAA
-1420          -1400          -1380
AATATTAATACAAATTTTAAACAGGGAAAAATCAAGAGAAGCAAAACAATGGAAAAAA
-1360          -1340          -1320
AATTGTTCAAATTTGTAGGTTGTAATTTTTTTGTTGCAATATTTAGCTTTATTTAAATT
-1300          -1280          -1260
GTTTTATCTTTCAATTTTGTGACTAAAATATTATTTAATAAATAAGCTGTTTAAATAA
-1240          -1220          -1200
TCTGTTTTGTTTAAATGCACCAAATACATTGCCTATATTTCACTGAGAGATGGAGAAAAA
-1180          -1160          -1140
TATTCATTTTCAAATGGGCTGTACTCGATTACTCTGAGCACTGTAATTATCTGTTTTTC
-1120          -1100          -1080
TAGGTACTGTAAAATAATGCTAATCTTAAACAATTTTGAACAAGAAGGAGCAAAACAAG
-1060          -1040          -1020
CCGACAGACTCACAGGATATCAAATTAGTCTAAAGAAAGAAGGGAATGCACAACAAACTC
-1000          -980          -960
AAGGGGACAAAACAGTGACTGATGAATTAGACAAGAAAAAGAGCAAGGAGCGCTCAA
-940          -920          -900
GATTGTTTAGCTATTTTGGTCACCCACAGCTGTTCTTATGCCTGCCTTCCCAAAAAAAA
-880          -860          -840
ACTGCTTAAGCCTCAAATTTCTCTCATGAGGGTCCAACATCAACCACTCAGAGGGCTG
-820          -800          -780
TAGTGTGCTGACCATCTAAAAACTGGGAAAAAGGGGTAATTACGTGCTTGTCCACAGGGC
-760          -740          -720
AGCTTCCCACAAATGGCACCTCACAGTCACTGAAGTGACCGGGTGAGGTCATAGGTCGAT
-700          -680          -660
CGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTGAATATGGGCATGGCCATGTGGAC
-640          -620          -600
GGGTGTGGTGGGGGCCTTGAACCGAAATCTTACAGCATCACTATACTAGGAAAAGCATT
-580          -560          -540
AAAACCTATTTTCGATTAGGGCTGATTTGAAATAAGGGTTAAGACACCAGAACGTCCTCT
-520          -500          -480
TATATATCAGCCAGGGGCGTGAACAAATATGAACAACATAATCATTTGGCTCAAAAAATCT
-460          -440          -420
CTGGATTGAAATCCATCAGGATCTATCACTGCAACCCCTCCCCATCCAACAAGTAATCCTG

```

圖 1A

-400 -380 -360  
| | |  
CAAGATAAGCCAGTATTGATCTGCTGCTAATCTAACTTTAGGGGAGGGAGTTGCGTCCCC  
-340 -320 -300  
| | |  
TTTAGACTCTGTGGCTACAGCTCATTCAATTTCAAATTGAGTTATGTGATTGTATGAAGCT  
-280 -260 -240  
| | |  
CAAACAGTCCCTTACGTCCCCATGTCCTTATTAGTCAACGCGAGACATGCAGGCCGCTGC  
-220 -200 -180  
| | |  
CATCAGTATCAGATTCATCCCATTCCAAGACTCCAATAGCTATTTCTGAGCACTGTAAGA  
-160 -140 -120  
| | |  
TGATAGTACATCCAGCCGGTGTCCCTCCATCACTTTCCCCCTACCTCATAGTTTTGCCT  
-100 -80 -60  
| | |  
CTTTCTCTCTCGGTCTGCTATTTCCCAAACCTCACTTAAGGTTGGGTCTATAATTAGCAA  
-40 -20 1  
| | |  
GGGGCCTTCGTCAGTATATAAGCCCCCTCAAGTACAGGACACTACGCGGCTTCAGACTTCT  
20 40  
| |  
CTTCTTGATCTTCTTAGACTTCACACATACCGTCTCGACATG

圖 1B

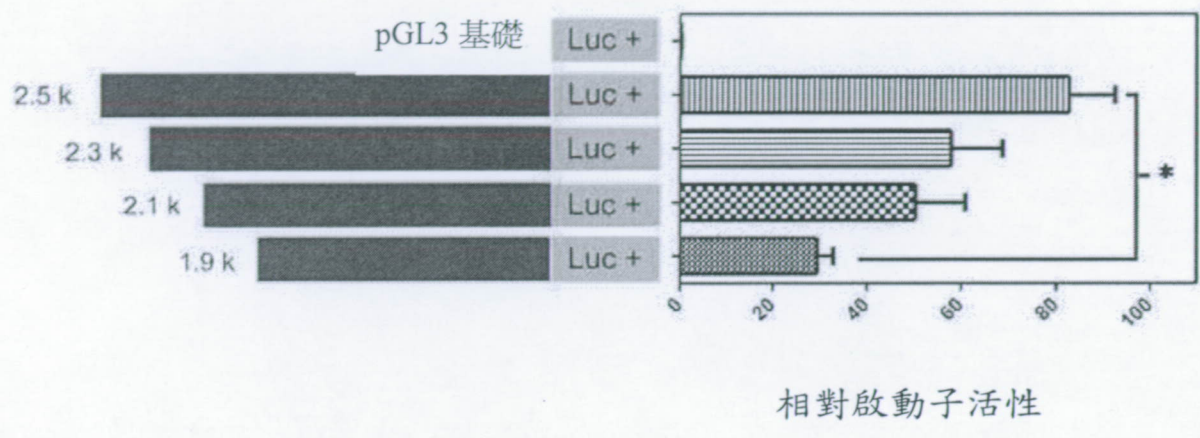


圖 2



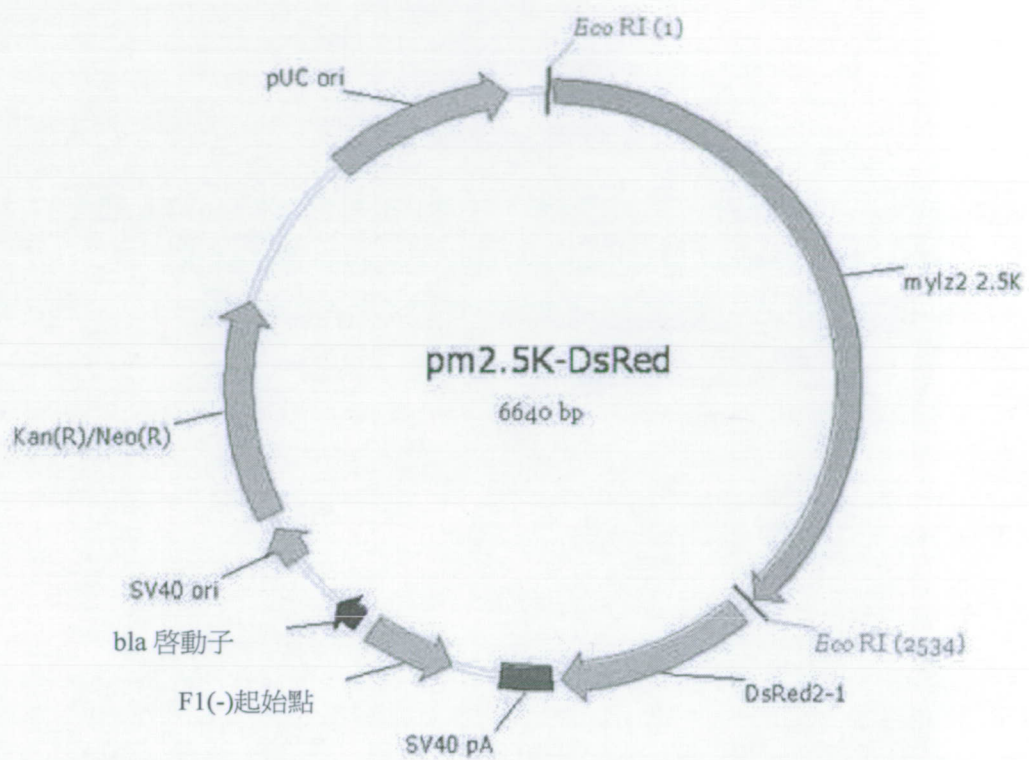


圖 3

1 M A S S E N V I T E F M R F K V R M E G  
1 atggcctcctccgagaacgtcatcaccgagttcatgcgcttcaaggtgcgcatggagggc  
21 T V N G H E F E I E G E G E G R P Y E G  
61 accgtgaacggccacgagttcgagatcgagggcgagggcgagggccgccctacgagggc  
41 H N T V K L K V T K G G P L P F A W D I  
121 cacaacaccgtgaagctgaaggtgaccaagggcggccccctgcccttcgcctgggacatc  
61 L S P Q F Q Y G S K V Y V K H P A D I P  
181 ctgtccccccagttccagtagggctccaaggtgtacgtgaagcaccgccgacatcccc  
81 D Y K K L S F P E G F K W E R V M N F E  
241 gactacaagaagctgtccttccccgagggccttcaagtgggagcgcgtgatgaacttcgag  
101 D G G V A T V T Q D S S L Q D G C F I Y  
301 gacggcggcgtggcgaccgtgaccaggactcctccctgcaggacggctgcttcatctac  
121 K V K F I G V N F P S D G P V M Q K K T  
361 aaggtgaagttcatcggcgtgaacttccccctccgacggccccgtgatgcagaagaagacc  
141 M G W E A S T E R L Y P R D G V L K G E  
421 atgggctgggaggcctccaccgagcgcctgtacccccgcgacggcgtgctgaagggcgag  
161 T H K A L K L K D G G H Y L V E F K S I  
481 accacaaggccctgaagctgaaggacggcggccactacctggtggagtcaagtccatc  
181 Y M A K K P V Q L P G Y Y Y V D A K L D  
541 tacatggccaagaagcccgtgcagctgcccggctactactacgtggacgccaagctggac  
201 I T S H N E D Y T I V E Q Y E R T E G R  
601 atcacctcccacaacgaggactacaccatcgtggagcagtagcagcgcaccgagggccgc  
221 H H L F L \*  
661 caccacctgttctctgtag

6640 GAAATCGCCC TTCGCTCGAG ATGCTGTGAA GTAATCTCTA CTTATCTTAA TTCTCTAAAA ATAATGCAAA AGTATTTCTTT AAAATGTTAA ATGAAGGGAT  
 100 ATTTTCAAAG CCAATCGACT TTTAGAAATAG ATAATGTTTTA ATTGACTAAC AGCAATTCAG CTACAAATTCAG AGTAAATGTTT TGAAGATAGA TATGAGTTAA  
 200 CAAAATCTCTG GACAACTACA TACTAATTTGA TGATTTAGTTT AAACCTATTA TTTTACATTA CAGGGAGTAT ACTACAATAT AAATACGATA AAACCTTGTGA  
 300 GCATGTATTA AAGCAATTA TAATTCCTCA ATCAAAATATC ACCCATATG TCCCATGCTC AGCATATATG AGTACACCCC TCTAGCCTGA TCTCAAGAGA  
 400 AAACGGGAGT APTTTATGTT TTGCCAGTTT AGTTCAGTCA TTGGGGGATG AGCAAAATCAT ACTAAAATGT ACGAATTAGT TCGTACGAAAT TAGCCACTAA  
 500 ATGAAAAAAGT TACAAAAATTG TGTGACCCC TCACAAAATCT GTCTTCATTC ATTTTCTTTT AGGCTTAGTC CCTTTATTAA TCAGGATTCG CCACAGAGGA  
 600 ATGAGCCACC AACTCATCCA GTGTATTTTT TATGCAGCGG ATGCCCATCC AGTTGCAGCC CTACACTGGG AAACACCCAA ATCTGTCTTT TAAATTAATA  
 700 TTTTAAATAG GAAGCTATAC AATGTTATAT TTGTGCATAT ACATTAGATT AGTCAGTACA CTCTCAGAAA TAAAGGTGCG CAAGCTGTCA CTGCGGTGGT  
 800 ACCTTTTCAA AAGGTACATA TTTGTACTTA AAGGTTCCAT ATTGATACCT CAAAAAGTATA TACTAGTGCC TAAAAATTTT AAGAGAAAACA CTTTTGTACT  
 900 TTTTAGGTAC TAATATATAC CTTGAGGTT TTAATATGGA CTTTAAAGGT ACAAAATTTT ATCAGTTGGA AAGGTACCAC CCCAGTGACA GTTTACACCC  
 1000 TTTATTTCTG AGAGTGAAGC TAAATCTGGA GCTTATTTAA CAAAATTAAT TATGATTAAC GTCCAAAAAC TACTACACCC ATATTTATAT GTTATTGAAA  
 1100 AATATTAAT ACAAAATTTA AAACAGGGAA AAATCAAGAG AAGCAAAAACA ATGGAATAAA AATTGTAGG TTTGTAATTTT TTTGTTGCCAA  
 1200 TATTTAGCTT TATTTAAAT TTTTATCTT TCAATTTTGT GACTAAAATA TTTATTTTAAAT AAATATAGCT GTTTAATAAA TCTGTTTTGT TTAATATGCAC  
 1300 CAAAATACAT TGCCTATATT CACTGAGAGA TGGGGAAAAA TATTCATTTT CAAAATGGGC TGTACTCGAT TACTCTGAGC ACTGTAATTA TCTGTTTTTT  
 1400 TAGGTACTGT AAAATAATGC TAATCTTAAA CAATTTTGGG ACAAGAAGGA GCAAAAACAAG CCGACAGACT CACAGGATAT CAAAATTAGTCAA TAAAGAAAAGA  
 1500 AGGGAATGCA CAACAACTC AAGGGGGACA AAACAGTGAC TGATGAATTA GACAAGAAA AGAGAGCAAG GAGCGCTCAA GATTGTTTAG CTATTTTGGT  
 1600 CACCCACAGC TGTTCCCTTAT GCCTGCCTTC CCAAAAAAAA CTGTCCTTAAG CCTCAATTTT CTCTTCATGA GGGTCCAACA TCAACCACTC AGAGGGCTGT  
 1700 AGTGTGCTGA CCATCTAAAA ACTGGGAAA AGGGGTAAT ACCTGCTTGT CCACAGGGCA GCCTCCACA AATGGCACCT CACAGTCACT GAAGTGACCCG  
 1800 GGTGAGSTCA TAGGTCGATC GGCAGAGAGA GAGAGAGTGC TGAATATGGG CATGGCCATG TGGACGGGTG TGGTGGGGG ACCTTGAACCG  
 1900 AAATCTTACA GCATCACTAT ACTAGGAAA GCATTAATAAC CTATTTTCGA TTAGGGCTGA TTTGAAAATA GGGTTAAGAC ACCAGAACGT CCTCTTATAT  
 2000 ATCAGCCAGG GCGGTGAATA AATATGAACA ACATATCAT TGGCTCAAAA AATCTCTGGA TTGAAAATCCA TCAGGATCTA TCACCTGCAAC CCTCCCCATC  
 2100 CAACAAGTAA TCCTGCAAGA TAAGCCAGTA TTGATCTGCT GCTAATCTAA CTTTAGGGGA GGGAGTTGCG TCCCCTTTAG ACTCTGTGGC TACAGTCAAT  
 2200 TCAATTTCAA TTGAGTTATG TGATTTGATG AAGCTCAAAC AGTCCCCTTAC GTCCCCATGT CCTTATTAGT CAACGGCAGA CATGCAGGCC GCTGCCATCA  
 2300 GTATCAGATT CATCCCATC CAAGACTCCA ATAGCTATTT CTGAGCACTG TAAGATGATA GTACATCCCA GCCGGTGTCC CTCCATCACT TTCCCCCTAC

圖 5A

2400 CTCATAGTTT TGCCCTCTTT TCTCTCGGTC TGCTATTTCC CAAACCTCAC TTAAGGTTGG GTCTATAATC AGCAAGGGGC CTTTCGTCAGT ATATAAGCCC  
2500 CTCAAGTACA GGACACTACG CTAGCCGAAG GCGGAATTCT GCAGTCGACG GTACCCGGGG CCGGGATCC ACCGGTCGCC ACCATGGCCT CCTCCGAGAA  
2600 CGTCATCACC GAGTTCATGC GCTTCAAGGT GCGCATGGAG GGCACCGTGA ACGGCCACGA GTTCGAGATC GAGGGCGAGG GCGAGGGCCG CCCCACGGAG  
2700 GGCCACAACA CCGTGAAGCT GAAGGTGACC AAGGGCGGCC CCCTGCCCTT CGCCTGGGAC ATCCTGTCCC CCCAGTTCCA GTACGGCTCC AAGGTGTACG  
2800 TGAAGCACCC CGCCGACATC CCCGACTACA AGAAGCTGTC CTTCCCCGAG GGCTTCAAGT GGGAGCGCGT GATGAACTTC GAGGACGGCG GCGTGGCGAC  
2900 CGTGACCCAG GACTCCTCCC TGCAGGACGG CTGCTTCATC TACAAGGTGA AGTTCATCGG CGTGAACTTC CCTCCGACG GCCCCGTGAT GCAGAAGAAG  
3000 ACCATGGGCT GGGAGGCCCTC CACCGAGCGC CTGTACCCCC GCGACGGCGT GCTGAAGGGC GAGACCCACA AGGCCCTGAA GCTGAAGGAC GCGGGCCACT  
3100 ACCTGGTGGG GTTCAAGTCC ATCTACATGG CCAAGAAGCC CGTGCAGCTG CCCGGCTACT ACTACGTGGA CGCCAAGCTG GACATCACCT CCCACAACGA  
3200 GGACTACACC ATCGTGGAGC AGTACGAGCG CACCGAGGGC CGCCACCACC TGTTCCTGTA G

圖 5B