

中華民國第 63 屆中小學科學展覽會 作品說明書

高中組 動物與醫學科

第三名

052012

解除 CRISPR(Cas9)辨認 PAM 序列的限制

學校名稱：臺北市私立復興實驗高級中學

作者： 高一 王譽恩 高一 高芮妮	指導老師： 馬瑪宣 呂毓蘋
---------------------------------	-----------------------------

關鍵詞：CRISPR/Cas9、DNA 重組技術(DNA recombination)、PAM 序列(PAM sequence)

摘要

CRISPR/Cas9(Clustered Regularly Interspaced Short Palindromic Repeat/CRISPR associated protein9)是合成生物學中重要技術，需引導 RNA 辨認特定的 PAM 序列-NGG 引導合成轉錄因子啟動或抑制下游基因。然而這個特定的 PAM 序列卻會因此限制了 crispr 技術可篩選的基因位點，造成限制。因此我們研發出 crisprTFv2，以 NAG 為例突破 PNGG 的限制，並透過定序得知更改後的 PAM 相互作用區序列為何，最後再挑選 crisprTFv2 可能辨認的其他 PAM 序列，未來再將此技術應用於探討參與 DNA 損傷修復的基因組合。

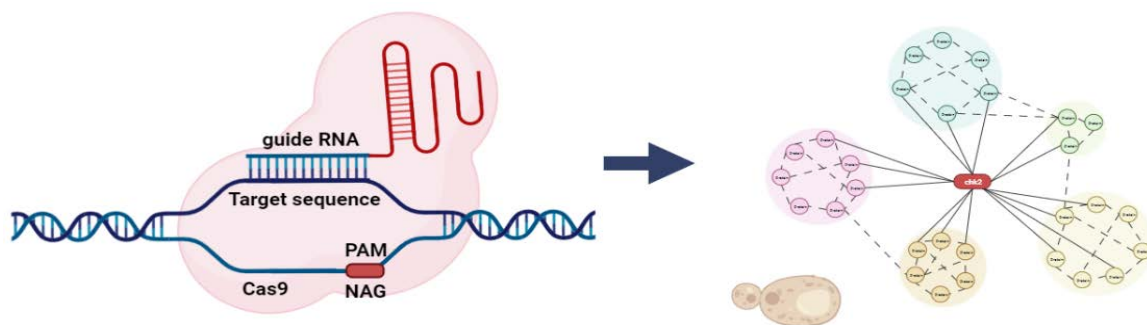
壹、前言

一、前言

(一)研究動機

在現今與基因相關的研究中，科學家們可以輕易的利用 CRISPR 基因編輯技術強制表現或抑制特定基因。然而，CRISPR 技術仍有其限制，例如:CRISPR/Cas9 的 Cas9 核酸酶是透過辨認 DNA 序列中一段三核苷酸 PAM 序列(Protospacer Adjacent Motif Sequence)鎖定剪切的目標，而這段三核苷酸序列必須是富含鳥嘌呤(Guanine-rich)的部位，此限制會讓 Cas9 核酸酶在 DNA 中的切點選擇減少。

現今使用的 CRISPR/Cas9 僅能辨認到目標 DNA 為 NGG(N 為任一個核苷酸)的 PAM 序列，本研究希望能利用酵母菌(Yeast)及大腸桿菌(*E.coli*)進行 CRISPR/Cas9 的相關研究，透過更換 Cas9 核酸酶上負責辨認 PAM 序列的 PI domain 之胺基酸序列，並以解鎖 PAM 序列-NAG 的限制為例，希望之後能利用此研究的方法解鎖其他 62 種 DNA 序列的可能，使此技術的使用更加廣泛與便利(圖一)。



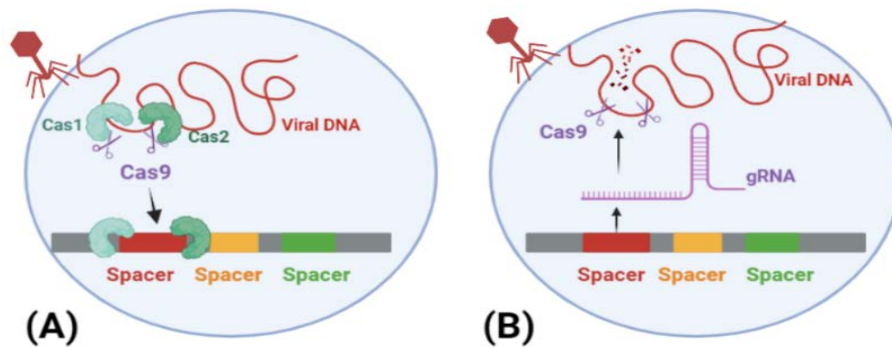
圖一: 本研究動機。CRISPR 受 PAM 序列的限制，本研究希望研發能辨認 PAM 序列 NAG 的 Cas9 核酸酶，應用在疾病基因表現的研究。

(二)研究背景

1.原始的 CRISPR (Clustered, Regularly Interspaced, Short Palindromic Repeats)

病毒自古至今便是各種細菌與動植物的剋星，為了能有效的對抗病毒侵擾，許多原核生物擁有一套後天免疫的機制，稱為 CRISPR 系統，使他們能記憶專一的病原，將攻擊自己的病原體 DNA 放入基因庫中，當下一次再遭到同一種病毒攻擊時，便能快速表現 RNA，有效應對。

CRISPR 在原核生物中是一串很長的基因序列，最前端(5'端)有一串 cas 序列 (又稱 cas cassette)，其後為一段"leader"序列以及 CRISPR 的重複序列。重複序列主要由兩部分重複排列而成，前面叫"repeat (R)"，後面是"spacer (S)"，其中，每個 RS 的組合裡 R 的部分皆為相同序列，S 為從病原身上抓下的 DNA 序列。每當有新的病毒來犯，CRISPR 機制中名為 cas9 的剪切蛋白會剪開病毒的一小段 DNA，以它為模板製作一段互補的 gRNA(guide RNA)，使細菌對此病毒存有記憶。當同一病毒再次入侵，CRISPR 透過複合體中的 gRNA 序列快速的辨認目標 DNA 的其中一股 protospacer 序列，在互補股中的 PAM 序列(Protospacer adjacent motif)也可輔助 Cas9 的切點辨認，最終精準剪切病原體 DNA，使細胞不被再次病毒入侵(圖二)。(Adli, M., 2018)

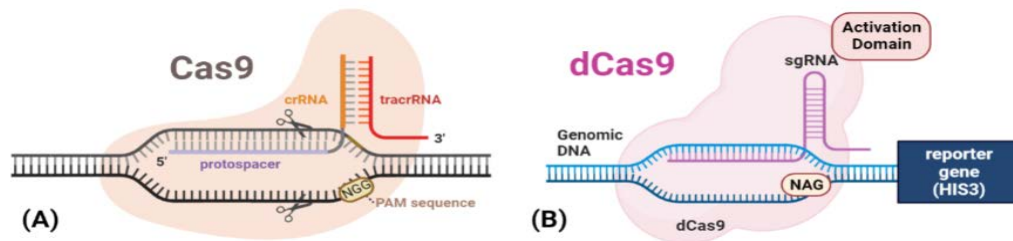


圖二: 細菌的 CRISPR 系統。(A)病毒第一次入侵細菌。在 Cas1 酶和 Cas2 酶辨認病毒的 DNA 後，Cas9 酶會將病毒的基因片段放入 CRISPR 序列中。(B)當病毒二度侵入細菌，細菌可以利用 CRISPR 序列的資料製作 gRNA，在最短時間內破壞病毒的 DNA。

2.CrisprTF(Clustered, Regularly Interspaced, Short Palindromic Repeats Transcription Factor)

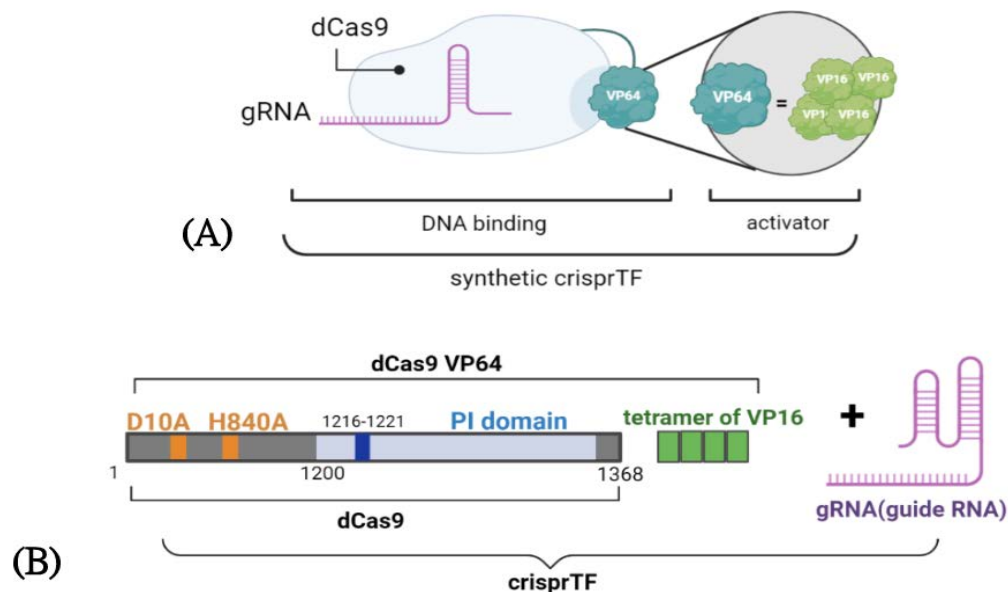
為了研究 Cas9 辨認特定的 PAM 序列，本實驗利用 CrisprTF 系統來研究，此系統含有 dCas9 (dead-CRISPR-associated protein 9), VP64 和轉錄因子。dCas9 這個詞是指“死”的 Cas9，它失去切割 DNA 的能力。Cas9 蛋白有兩個 DNA 核酸內切酶結構域，HNH (或 McrA) 和 RuvC 結構域，每個結構域都賦予對目標雙鏈 DNA 的一條鏈的切割活性。這些結構域中催化殘基的胺基酸取代，HNH 結構域的 D10A 和 RuvC 樣結構域的 H840A，消除了 DNA 切割活性，同時保留了其特定的 DNA 結合活性 (Park JJ et al., 2017; Daan J.W. Brocken et al., 2018)。

PAM 序列是一段簡短的序列，通常為 2 至 5 個鹼基對，位於目標 DNA 旁邊，需要識別「自我」與「非自我」。野生型 CRISPR 的 PAM 序列為 NGG，而此研究中的 crispTF v2 所能辨認的 PAM 序列則為 NAG(N 可為 A、T、C、G 任意核苷酸)。PAM 序列對於 CRISPR 系統十分重要，因為它能代表剪切酶的切點位置。(Gleditsch D et al., 2019)我們主要改變的是 dCas9 酶中 PI 域上的序列 1216-1221，它決定了 dCas9 中 PAM 序列的識別(圖三)。



圖三: Crispr/Cas9 與 dCas9 機制之比較圖。(A) 原始 CRISPR/Cas9。(B) 修改的 CRISPR/Cas9，稱為“CRISPR/dCas9”。它缺乏核酸內切酶活性，但當 dCas9 識別正確的序列時，標記為藍色的 reporter 基因會表達組氨酸(HIS3)基因。因此，細胞可以自行製造 Histidine 營養，並允許細胞在 SCM-His 的培養基上生長。

VP64 是一種轉錄激活因子，由四個相同且串聯再一起的 VP16 (單純皰疹病毒蛋白 16) 與甘氨酸-絲氨酸(GS)接頭連接而成。VP64 可將轉錄機制召集到特定序列旁，

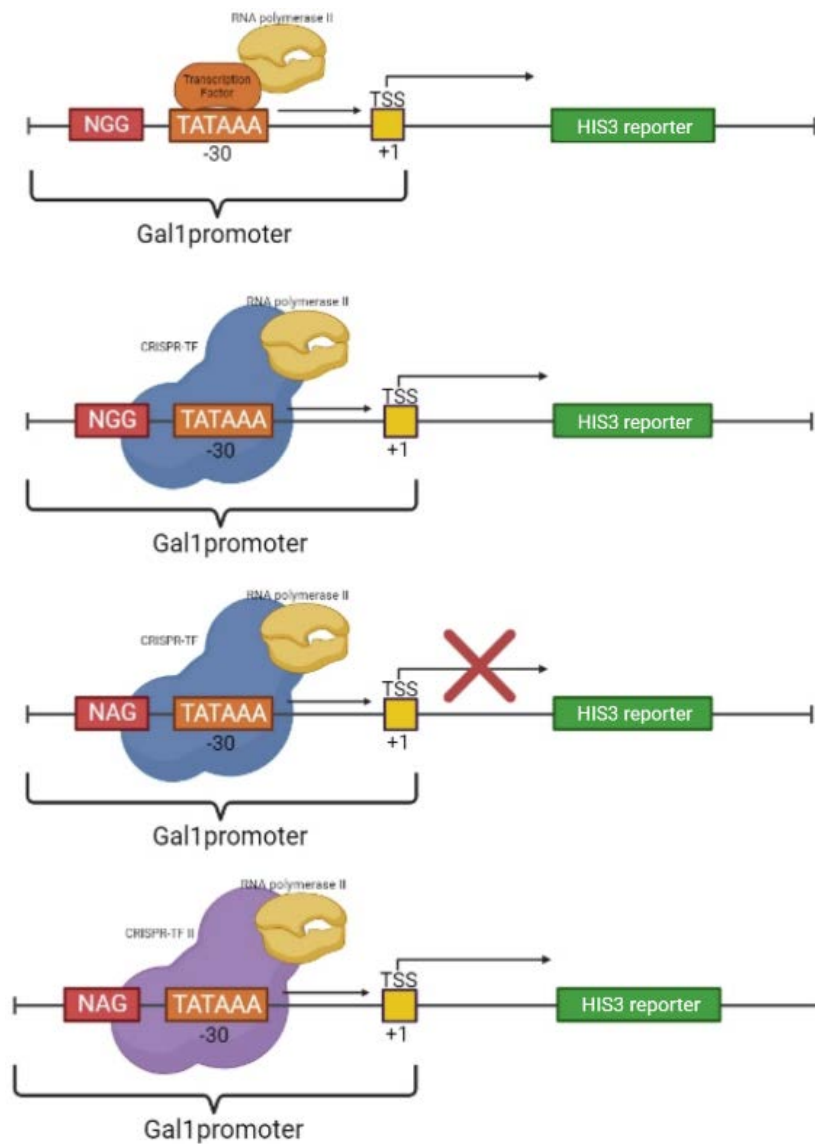


從而引起靶向基因調控。(Gasiunas G, 2012;Jinek M, 2012)(圖四)

圖四：此實驗使用的 crisprTF 之結構組成以及帶有的序列。(A) crisprTF 由 dcas9 和 VP64 組成，而 VP64 又是由 4 個相同的 VP16 組成。VP16 是轉錄因子，可啟動基因表現的調控。(B) crisprTF 的序列組成，可以和野生型 Cas9 結合的序列結合，但差別在於其無法剪切序列。

基因表現過程中，轉錄作用時，RNA 聚合酶 (RNA polymerase) 與啟動子 (promoter) 的結合能力會影響基因表現的程度，藉此對基因的轉錄做出正向或是負向的調節。而在啟動子中，當特殊的結合蛋白質 (轉錄因子，Transcription Factor, TF) 與基因轉錄作用起始位置 (Transcription Start Site, TSS) 鄰近的序列 (調控序列，Transcription Factor Binding Site, TFBS, Regulation Site)，例如：TATA box、CAAT box、GC box 等，進行交互作用後，能使 RNA 聚合酶與 DNA 模板結合並轉錄下游基因。

以 TATA box 為例，此調控序列常見於高等真核生物中，通常坐落於距離 TSS 位置前 30 個核苷酸的位置，其序列以 A、T 為主，共有序列的型式為 TATAAAA，故稱之 TATA box。在 TATA box 與轉錄因子結合後，能幫助 RNA 聚合酶將 DNA 展開雙股結構，以便 RNA 聚合酶沿著鑄模股合成 mRNA。(Wang Y, Jensen RC, Stumph WE., 1996)(圖五)



圖五:TATAbbox 與不同 crisprTF 結合，搭配不同 PAM 序列的結合情形。TATAbbox 與 crisprTF v1 或 crisprTF v2 結合後，啟動 HIS3 轉譯成 mRNA。crisprTF v1 能辨認 PAM-NGG 啟動 HIS3 基因表現;crisprTF v2 能辨認 PAM-NAG 啟動 HIS3 基因表現;crisprTF v1 則無法。

(三)研究目的

- 1.更換 Cas9 核酸酶上負責辨認 PAM 序列的 PI domain 上胺基酸序列
2. 以 N 序列產生 PI domain Library
3. 利用 Spotting Assays 篩選出 crisprTF v2
4. 透過大腸桿菌 DNA 轉型實驗(*E.coli* transformation)確認 crisprTF v2

5. 透過酵母菌 DNA 轉型實驗確認 crisprTF v2 的正確性
6. DNA 定序(Sanger Sequencing)
7. 挑選 crisprTF v2 可能辨認的 PAM 序列

貳、研究設備及器材

一、細胞株:

(一) 酵母菌(*m555*)

(二) 大腸桿菌(*Escherichia coli*)

二、藥品及試劑

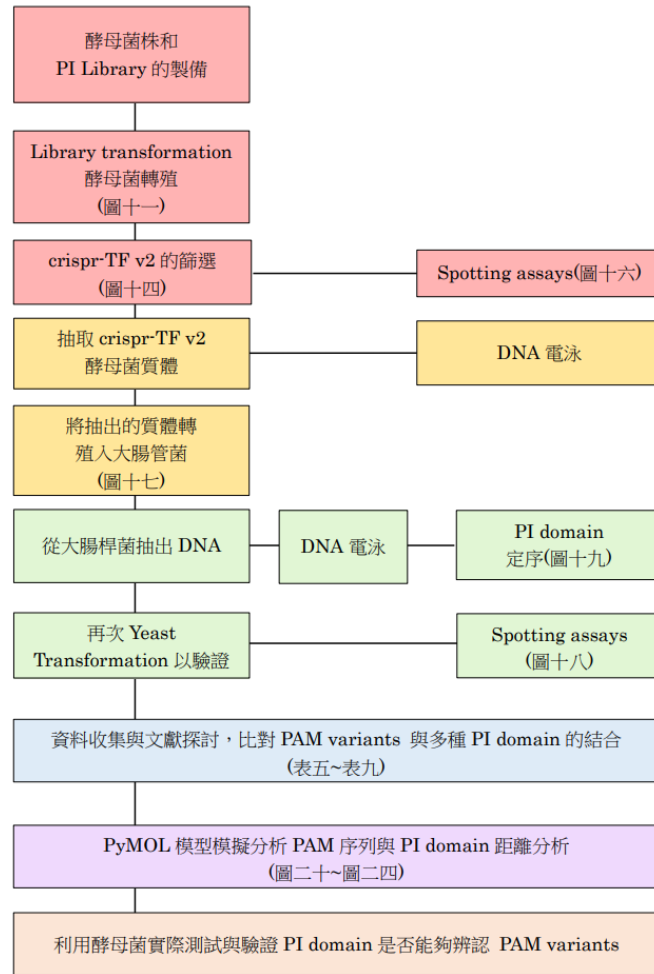
溴化乙錠(EtBr)、ddH₂O、1M and 0.1M 乙酸鋰 LiAc(Lithium acetates)、50%(PEG)、YPD broth、SCM -Ura YPD 培養基、LB (Lysogeny broth)、LB + Ampicillin broth、EtOH、DNA Loading Dye、DNA Ladder、Plasmid DNA Miniprep Kit、10% SDS、5M NaCl、1M Tris、PH8.0 0.25M EDTA、Triton X-100、Dox(doxycycline)、3AT(3-amino-1,2,4-triazole)

三、器材

恆溫水槽、微量吸管分注器(pipette)、離心機、微量吸管尖(pipette tip)、96孔 PCR 反應盤、微波爐、電泳照膠系統、電腦、分光光度計、電泳槽、30°C與 37°C 恆溫箱、Spotter、eppendorf、-80°C 冰箱、4°C 冰箱。

參、研究過程或方法

一、研究架構圖



圖六：實驗流程

二、研究方法及步驟

(一)細胞培養(cell culture)

1.酵母菌培養:

將酵母菌培養於 YPD 中，並放置於 30 度的培養箱隔夜培養。在實驗前 recover，將菌液離心，加入新的 YPD 再培養 4 至 4.5 小時。

2.大腸桿菌培養:

將大腸桿菌培養於約 4mL LB 溶液中，並放置於 37 度培養箱培養至少 8 小時。

3.細胞繼代培養:

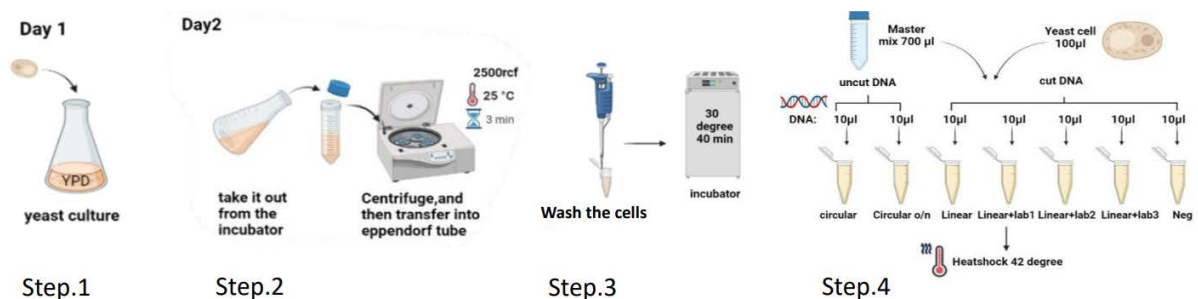
由於酵母菌複製 DNA 會有極限(copy number)，複製的效率會隨時間衰減。因此我們每個禮拜會將原本生長在培養基上的菌落挑六顆畫在新的盤子上，使其繼續生長，並保持在最適合做實驗的狀態。

(二)YPD 培養基的製備

將 5 g 瓊脂、5.5 g 葡萄糖和 175 mL ddH₂O 添加到可高壓滅菌的燒瓶中。接著加入 2.05 g SCM 混合併加入 75mL ddH₂O 到另一個可高壓滅菌的燒瓶中。將兩個高壓滅菌瓶放入滅菌器中，之後，將不同的胺基酸加入帶有 SCM 混合液的高壓滅菌燒瓶中，充分混合並加入帶有瓊脂的高壓滅菌閃光燈中。趁熱將培養基倒入培養皿中。最後，等待瓊脂混合物冷卻凝固。

(三)Library DNA 轉殖實驗(Yeast Library DNA Transformation)

將液態培養後的酵母菌以 4000rpm, 25°C 離心 3 分鐘分離出酵母菌細胞。用 50ml 水清洗細胞，然後在 25°C 下以 5000 rpm 的速度離心 3 分鐘。用 1000 µl 0.1M LiAc 洗滌細胞，混勻。然後在 5000 rpm、3 分鐘、25°C 下旋轉細胞。該步驟重複兩次。加入 1000 µl LiAc，混勻。存放在 30°C 培養箱中 30 分鐘。準備 Master Mix，在 eppendorf 中加入 10µl DNA、700ul Master Mix 和 100µl 酵母細胞。放入 30°C 培養箱 30 分鐘後，再放入 42°C 恆溫水槽中 20min。最後，離心完使用 SCM 培養基塗盤進行培養。(每份 Master Mix 含 65µl ddH₂O, 36µl 1M LiAc, 5µl carrierDNA, 240µl 50% PEG)



圖七: LiAc yeast transformation 流程示意圖

(四)酵母菌 DNA 萃取(Yeast genomic DNA Extraction)

在實驗進行前，在冰上冷卻 EtOH。在 YPD 培養基中開始 3-5ml 過夜培養，以 5000g、1 分鐘、4°C 離心細胞準備 Lysis Buffer。將 200 µl 緩衝液混和混勻。每管加入 100µl 玻璃珠和 200ul 25:24:1 苯酚：氯仿：異戊醇，以 2600rpm(3min)離心，然後加入 200µl TE buffer，倒置混勻。以 18000g 4°C 5min 離心，將上清液轉移至新的 1.5 µl 試管中，加入 1ml 100% EtOH。在冰上靜置 15min，18000g 4°C 離心 15min。小心吸去上清，用 400 µl TE Buffer 混和混勻。加入 3µl 10mg/ml RNase，在 37°C 環境下等待 5min。加入 5.4 µl 7.5M 醋酸銨(ammonium acetate)和 1ml 100% EtOH，混和混勻。在冰上靜置 15min，以 18000g 離心 15min。小心吸去上清液，高速真空(60°C)乾燥沉澱至少 2 小時。用 30 µl TE 緩衝液混和均勻，於 4°C 或 -20°C 保存。(每份 Lysis Buffer 含 846µl ddH₂O, 100µl 10% SDS, 5M NaCl, 10µl 1M Tris PH=8.0, 4µl 0.25M EDTA, 20µl Triton X-100)

(五)大腸桿菌 DNA 轉殖(*E.coli* DNA Transformation)

實驗開始前，先在冰上解凍 *E.coli* competent cells (10 分鐘)。在每一個 1.5µl 試管中加入 2-10µl candidate DNA 以及適當的 *E.coli* competent cells。在冰上靜置細胞/DNA 混合物 10 分鐘，接著，將混合物放入 42°C 恆溫水槽中 42 秒。將 1.5µl 試管置於 25°C (室溫) 3 分鐘。每 1.5µl 試管中加入 1ml LB 培養液，在 37°C 恆溫箱培養 1 小時。將細胞以 15000 rpm 的速度離心 3 分鐘。用 100 µl LB，混和混勻，在 LB +Amp 培養基上進行培養。

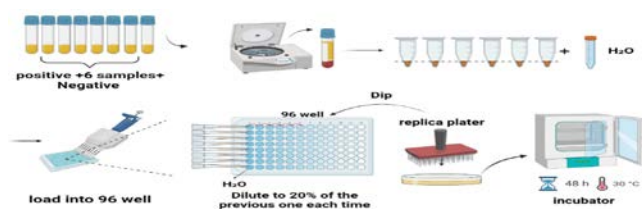
(六)大腸桿菌 DNA 萃取(*E.coli* genomic DNA Extraction)

在 LB 培養基中開始 3ml 過夜培養。以 15000g 離心 1 分鐘，接著加入 200 µl PD1 Buffer (添加了 RNase A)，完全離心使細胞沉澱。然後在樣品中加入 200 µl PD2 Buffer，然後輕輕顛倒管子 5 次混勻。加入 300 µl PD3 Buffer，顛倒 5 次混勻。15000g 離心 3min，重複兩次。將上清液轉移到 PDH 管中，15000g 離心 30 秒，丟棄廢液。將 PDH 管放回原處。加入 400ul W1 Buffer，以 15000g 離心 30 秒，重複兩次。15000g 離心 3min，烘乾管子。將乾燥的 PDH 管轉移到新的 1.5 µl 試管中，在管中加入 50ul Elution Buffer。等待 5 分鐘。15000 g 離心 5 分鐘。4°C 保存 DNA。

(七)生長定量分析(Spotting assay)

確定點樣試驗的稀釋比例後，每管移取 1mL 細胞，5000rpm 離心 1 分鐘，丟棄上清液使每管中的細胞量相等。(點樣試驗的比較只有在初始細胞數相同的情況下才有意義。)加入 200µL ddH₂O 混和均勻細胞。將細胞液轉移到微孔板的第一行孔洞並進行連續稀釋。(稀釋倍數越大，

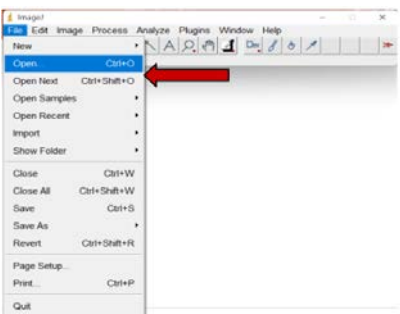
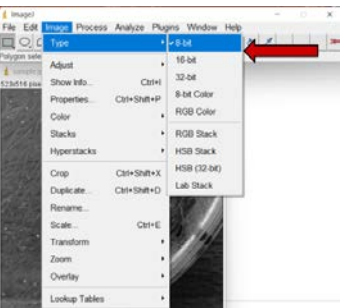
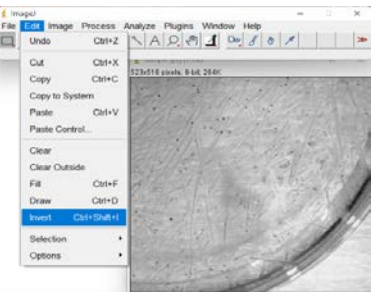
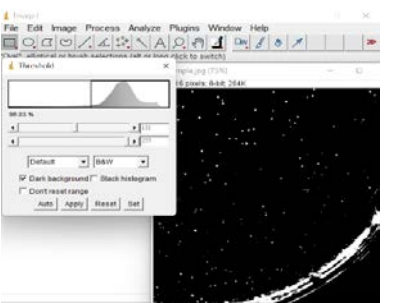
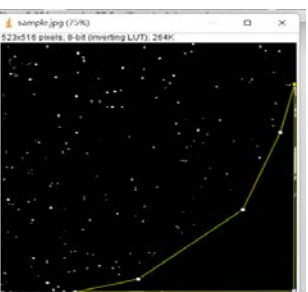
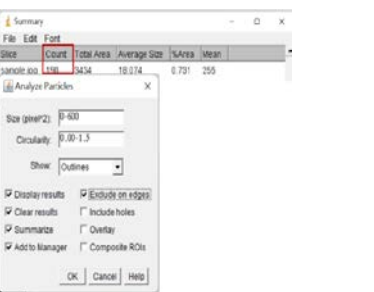
樣品差異越明顯。)用酒精清洗複製 96 孔盤。在不同培養基中以 spotter 蓋上細胞。待細胞乾透後，放入 30°C 培養箱中培養 3-5 天。



圖八: 生長定量分析過程

(七)細胞計算方法

如圖九所示，在此實驗中，我們使用 ImageJ 進行細胞數量的量化。

<p>圖 1: 打開 Image J 程式。</p> 	<p>圖 2: 將檔案轉換為 8-bit 形式。</p> 	<p>圖 3: 調整圖片使菌落更清楚。</p> 
<p>圖 4: 調整臨界值消除干擾。</p> 	<p>圖 5: 減去會干擾計算的部分。</p> 	<p>圖 6: 計算總細胞數，統合結果。</p> 

圖九: 使用 ImageJ 計算菌落之方法與流程。

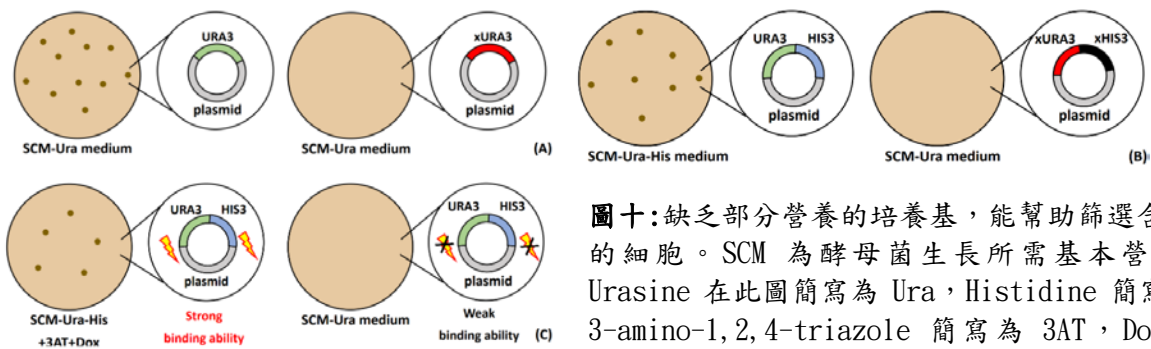
(八)利用培養基篩選細胞

由於本實驗使用的酵母菌菌種 m555 為 Urasine 與 Histidine 的營養缺乏型(auxotroph)，因此我們利用缺少部分營養的培養基來篩選出含有 crisprTF v2 的酵母菌細胞。

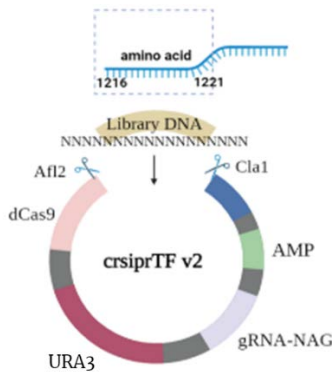
由於酵母菌需有經實驗送入帶有 URA3 基因的 plasmid 才有能力自行製作 Ura，因此圖十(A)中生長於 SCM-Ura 培養基的菌落都帶有 URA3 基因。圖(B)亦為同樣的原理，生長於 SCM-Ura-

His 培養基，菌落除了要有 URA3 基因外，也需帶有 HIS3 基因，能自行製造 Histidine。圖十(C)中，3AT(3-amino-1,2,4-triazole)是過氧化氫酶抑制劑，會與含有血紅素的活性四聚體形式的活性中心共價結合來抑制過氧化氫酶活性，因此能提升檢測的專一性，抑制生長能力較弱的菌落生長 (Mitsuyoshi Ueda et al., 2003)。Dox 在此則為誘導劑，能增加基因的表現量。(Taymaz-Nikerel H, et al., 2018)因此，酵母菌細胞的 plasmid 除了要帶有 URA3 與 HIS3 基因，其 PI domain 也要有足夠能力辨認 PAM 序列，才能啟動 HIS3 基因，成功在 SCM-Ura-His+3AT+Dox 培養基生長。

由(圖十)，實驗使用的 plasmid 帶有 URA3 基因，並以 HIS3 基因作為 dCas9 的 reporter gene，因此我們能利用此篩選方式挑選出帶有 crisprTF v2 的酵母菌細胞。



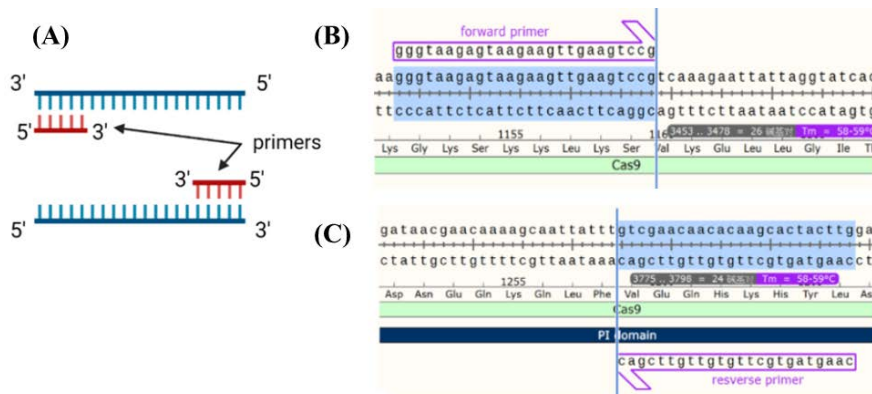
圖十: 缺乏部分營養的培養基，能幫助篩選含目標基因的細胞。SCM 為酵母菌生長所需基本營養成分，Urasine 在此圖簡寫為 Ura，Histidine 簡寫為 His，3-amino-1,2,4-triazole 簡寫為 3AT，Doxycycline



圖十一: crisprTF v2 的質體全貌。Afl2 與 Cla1 為來剪切質體的兩種限制酶，黃色為我們送入的隨機胺基酸序列(Library DNA)。

(九)引子(Primer)設計

引物是定序反應的起始信號。定序進行兩次，一次從頭開始，另一次從相反的方向 (reverse) 開始，使測序結果更加準確。



圖十二:(A)引子(primer)示意圖(B)實驗使用的兩種 primer。引子(primer)是短的單鏈 DNA 片段，引導待測序 DNA 片段的上游和下游區域。我們遵循引子設計原則自己找到了這兩個引子。在下面的實驗中，我們將使用這兩個引子進行測序。

1. 引子設計原則：

(1) 18-24 個鹼基對的長度。若引子太短，可能使定序不精確；反之，則可能使反應速率降低。

(2) 溶解溫度(Tm 值) 介於 50-60°C。Tm 值是一半引子(primer)將從 DNA 解離的溫度。適當的 Tm 值設置可以讓大部分引子與模板結合。

(3) G/C 含量 40-60%。

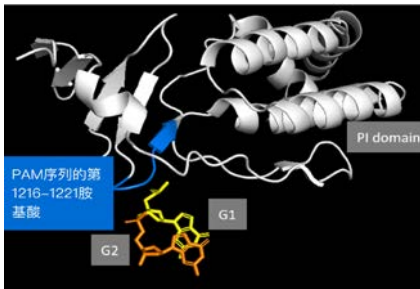
(4) 以 1-2 個 G/C 對開始和結束。G/C 對有 3 個氫鍵，比 A/T 對多一個，因此可以促進與模板的更強和更完整的結合。

(5) 位於目標基因前後至少 20 個鹼基對。前 100 個鹼基對的測序結果較不準確。因此我們必須在目標基因前後至少設計 20 個鹼基對的引子，使目標基因被完整正確的定序。

肆、研究結果

一、更換 Cas9 核酸酶上負責辨認 PAM 序列的 PI domain 上胺基酸序列

我們利用軟體「PyMOL」來模擬 crisprTF v2+gRNA(NAG)的結構，並發現 PI domain 第 1216-1221 序列區域與目標基因的 PAM 序列位置十分接近。因此推測此區可能是控制 PI domain 辨認的重要位置，因此，在實驗中選擇更改這部分的序列。



圖十三:白色區域為辨認 PAM 序列的 PI domain；藍色部份為 PAM 序列的第 1216-1221 胺基酸；黃色與橘色的 G1 和 G2 則分別為 PAM 序列(NGG)的第一個 G 和第二個 G。

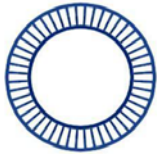
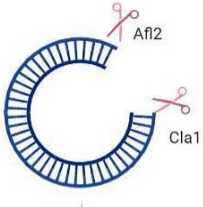

二、以 N 序列產生 PI domain Library

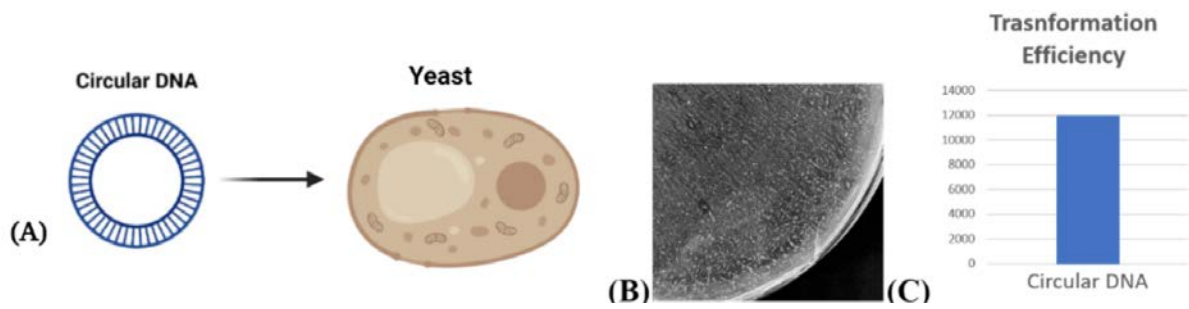
(一)實驗目的:

我們建立 PI domain 圖庫(PI domain Library)，使 6^{20} 種胺基酸序列組合的序列片段與載體接合，並篩選出含有能使 PI domain 辨認到新的 PAM 序列之重組 DNA 的酵母菌。

(二)實驗結果:

表一: Library Transformation 實驗所使用 DNA 種類

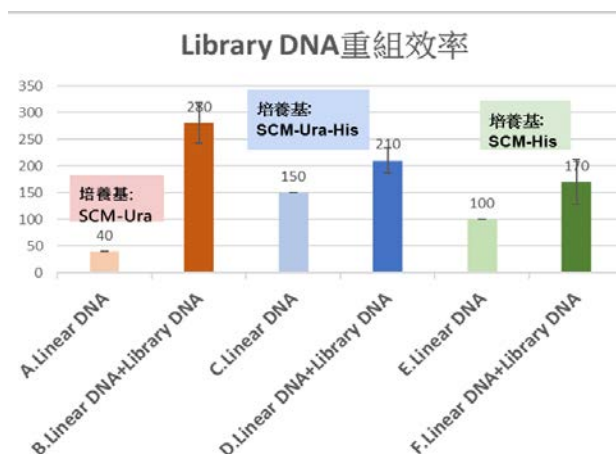
DNA 種類	circular DNA	Linear DNA	Library DNA
圖示			



圖十四: PI domain Library 之效率。(A)此步驟主要將質體(circular DNA)送入酵母菌中，並於 SCM-Ura 培養基中生長。(B)為四分之一的培養基生長出的菌量，(C)因此，我們可推算總共約有 12,000 個菌落生存，此數目也表示了此實驗能長出之菌落的最多數目。

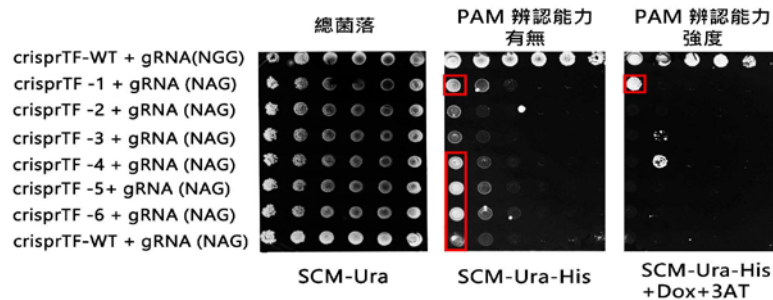
表二: Library 酵母菌轉型實驗結果。

培養基:“-”表示某營養缺陷；U 為 Ura；H 為 His			
質體種類	(1) -U	(2) -U-H	(3) -U-H+Dox+3AT
Linear DNA 			
Library 質體 			



圖十五: 紅色長條圖表示的數據為在 SCM-Ura 培養基的結果；藍色為 SCM-Ura-His 培養基；綠色為 SCM-His 培養基；利用長條圖比較菌落數目後，可發現 Linear DNA+Library DNA 組別的菌落數皆較 Linear DNA 多，此結果顯示擁有 LibraryDNA 的組別因能使 PI domain 辨認新的 PAM 序列(NAG)，也表示我們從龐大的 PI domain 圖庫中，獲得有用的序列。

三、利用 Spotting Assays 篩選出 crisprTF v2



圖十六: Spotting assay 的結果。

此實驗中，以 crisprTF-WT(Wildtype)+gRNA(NGG)為正控制組，由於其本身就辨認 PAM-NGG，因此搭配 gRNANGG，便容易啟動 CRISPR 系統；以 crisprTF-WT+gRNA(NAG)為負控制組，crisprTF-WT 只能辨認 PAM-NGG，因此當搭配 gRNA(NAG)時，就無法啟動 CRISPR 系統。crisprTF-1~6 為從 Library 酵母菌轉殖實驗中生長出的菌落。皆能在 SCM-Ura, SCM-Ura-His, SCM-Ura-His+Dox+3AT 之培養基生長，而紅框框起的部分為生長較密集的菌落，其可能具有較好的 PAM 序列辨認能力。(圖十六)

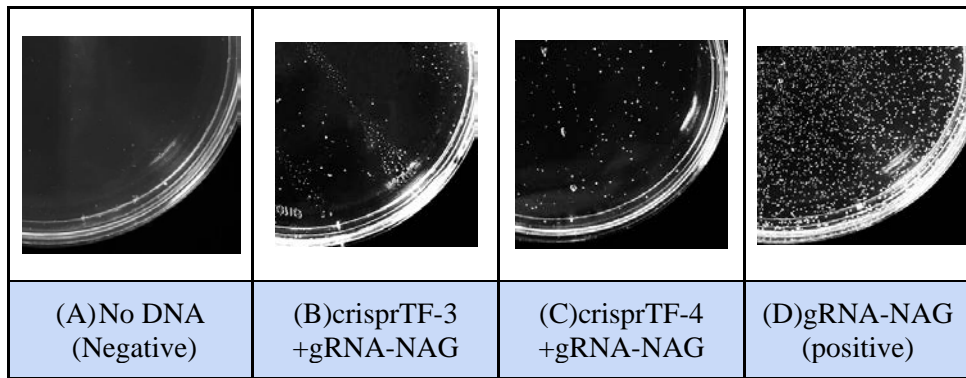
四、透過大腸桿菌 DNA 轉型實驗(*E.coli* transformation)確認 crisprTF v2

(一)實驗目的：

通過在 LB+Ampicillin 培養基上篩選具有質體的大腸桿菌細胞。另一方面，大腸桿菌的複製數(copy number)高於酵母菌。將我們的 candidate DNA 保留在大腸桿菌中可以增加質體的數量並促進後續的定序步驟。

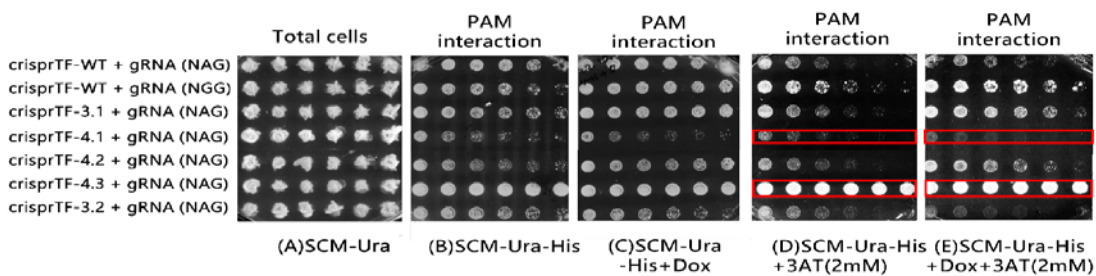
(二)實驗結果：

以下是我們利用 *E.coli* transformation 篩選帶有 crisprTF-3 及 crisprTF-4 的菌落。其他的 crisprTF v2 也是透過此方式，分批篩選出帶有質體的細胞。



圖十七: E.coli Transformation 結果。

五、透過生長定量分析(Spotting Assay)確認 crisprTF v2 的正確性



圖十八: 利用在 SCM-Ura 培養基繼代培養的細胞進行 Spotting assay 的結果。

第一行的 crisprTF-WT+gRNA(NAG)是負控制組；第二行 crisprTF-WT+gRNA(NGG)是正控制組，由於養菌時我們發現菌況不佳，在實驗結果只能從(A)(D)(E)組別中辨認正、負控制組。因此，我們只會針對這三組結果進行討論。為了進行強度的辨認，我們同樣在其中幾個培養基中加入 Dox(Doxorubicin)來提升 crisprTF 表現量；加 3AT(3-amino-1,2,4-triazole)作為競爭抑制劑，降低背景值。

六、DNA 定序(Sanger Sequencing)

經定序結果得知，我們成功更改 crisprTF-4.1 之 PI domain 第 1216-1221 序列，由 Ser,Ala,Gly,Glu,Leu,Gln 更改為 Trp,Cys,Gly,Cys,Ala,Gly。

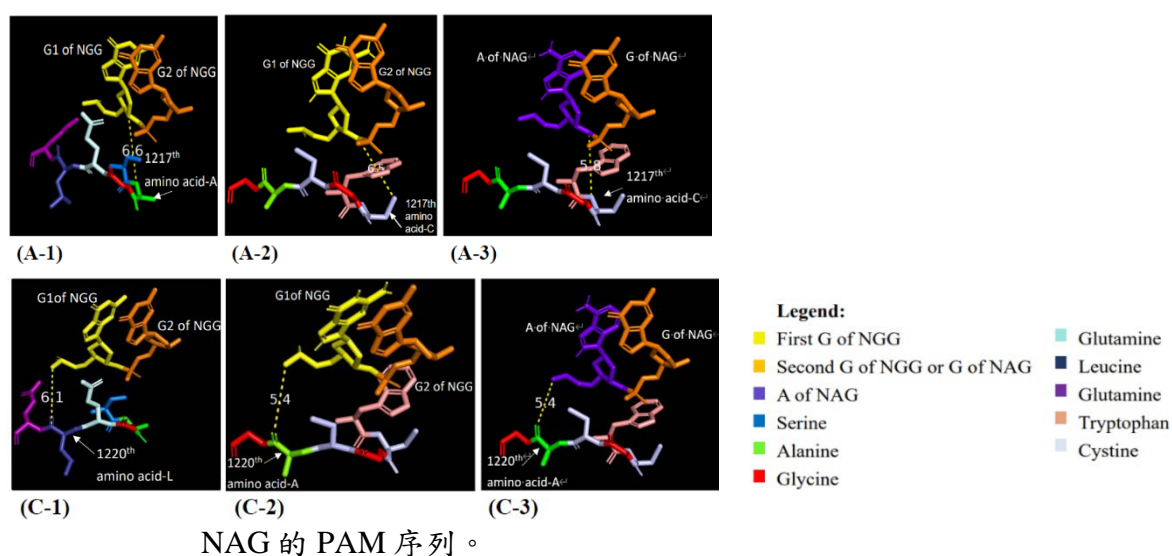


圖十九:酵母菌測序結果。原始 crisper-NGG 與 SAGELQ 結合。

dCas9-NAG(crisprTF-4.1) 與 WCGCAG 結合。(1216-1221)

七、模型模擬

分子間的距離越近，兩者有交互作用的機率越大。在定序後，我們觀察 crsiperTF v2 在 PyMOL 的模擬模型，並發現 PAM 序列與 PI domain 第 1216-1221 的距離在變換 6 個胺基酸序列後變短了。例如:第 1217 個胺基酸在從 Alanine 變為 Cysteine 後，距離從 6.6 Å 變為 5.8 Å，這個現象可解釋為何新的 crisperTF v2 能成功改為辨認

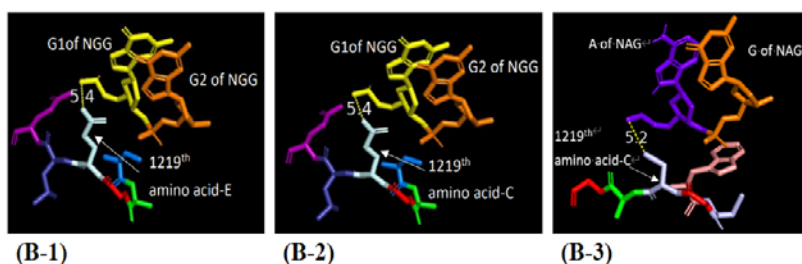


圖二十: 利用 PyMOL 程式模擬 dCas9 中

PI domain 的 1216-1221 序列與 PAM 的距

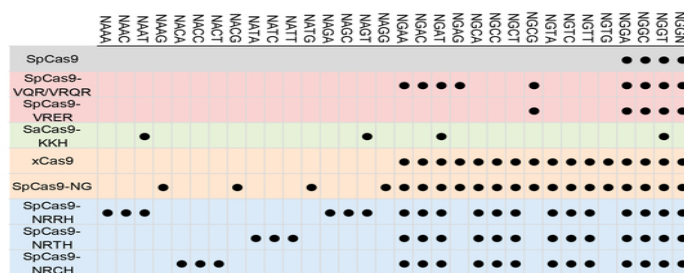
離關係(A-1)(B-1)(C-1)是原始的 crisperTF-

NGG ; (A-2)(B-2)(C-2)是 crisperTF-4.1-NGG ; (A-3)(B-3)(C-3) 是 crisperTF-4.1-NAG 。



八、挑選 crisperTF v2 可能辨認的其他 PAM 序列

(一)選擇以 NAG 為起點解鎖 PAM 序列限制



圖二十一:目前針對解除 PAM 序列限制的研究成果。(Miller SM, Wang T, et.al., 2020).

原始 crisprTF 辨認的 PAM 序列為 NGG，因此初期研究大部分以將 PAM 序列解鎖為 NGN 試圖解除 PAM 序列的第一個 G。接著，研究開始朝解鎖 PAM 序列的第二個 G 的方向，致力解鎖為 NNG。近期，則是欲解鎖其餘的可能性，打造能辨認 PAM-NNN 的 CRISPR。

	1216	1221
	胺基酸序列	能辨認的PAM序列
original crispr	SAGELQ	NGG
SpG[1]	SAKQLQ	NGN
xCas9[2]	SAGVLQ	NGN
SpCas9-NRRH[3]	SAGVLH	NAC NAA NAG NGN
SpCas9-NRTH[3]	SASVLH	NAT NGN
crisprTF-4.1	WCGCAG	NGG NAG ?

*紅字表示更改到的胺基酸序列。
 *N:A, T, C, G任意一種核苷酸。
 *R:A or G
 *H:A or C or T

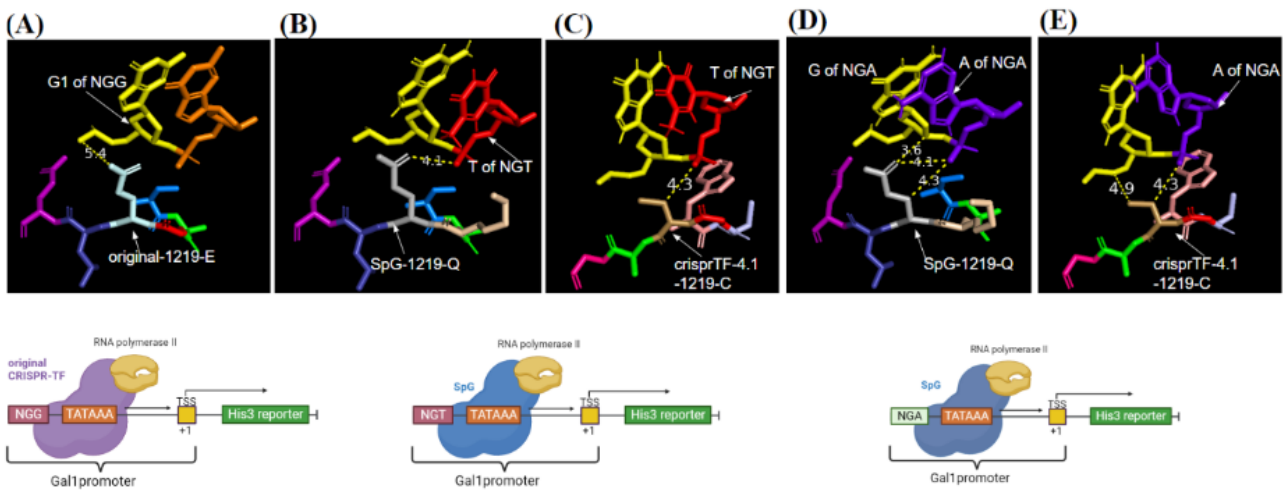
表五:original crisprTF 與 PAM variants、crisprTF-4.1 之胺基酸序列比較以及其能辨認的 PAM 序列。

(1) SpG

表六:SpG、crisprTF-4.1、originalTF 之 1216-1221 胺基酸比較。

	1216	1217	1218	1219	1220	1221
	Ser	Ala	Gly	Glu	Leu	Gln
original crisprTF (NGG)	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small	Polar Negative Charged Carboxylic	Non-polar Aliphatic Branched	Polar Nitrogen- containing Amides
SpG	Ser	Ala	Lys	Gln	Leu	Gln

(NGN)	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small	Polar Nitrogen- containing Amides	Non-polar Aliphatic Branched	Polar Nitrogen- containing Amides
TF4.1 (NAG)	Trp	Cys	Gly	Cys	Ala	Gly
	Polar Aromatic Phenylic Nitrogen -containing	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small



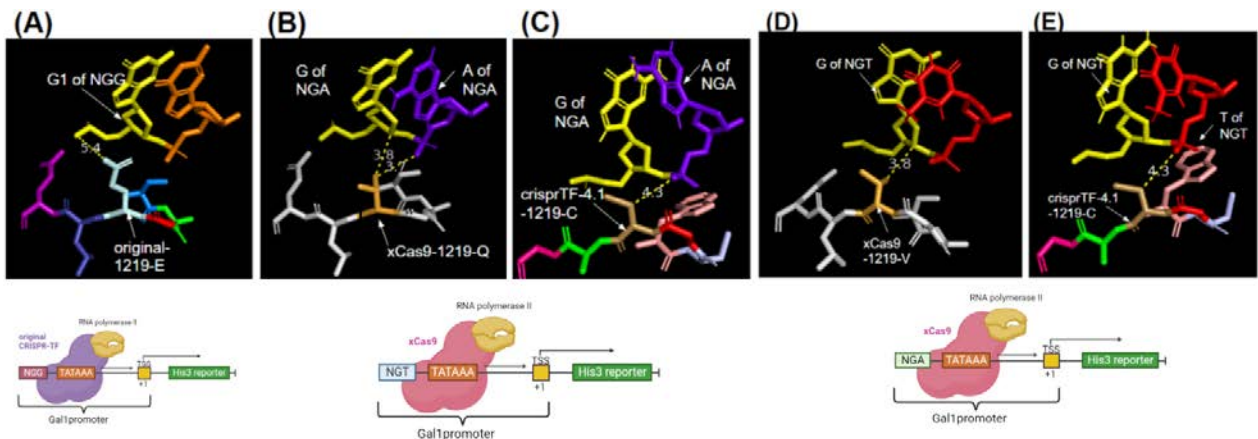
圖二十二:(A)原始 crisprTF 與 PAM-NGG 之相對位置。(B)SpG 與 PAM-NGT 之相對位置。(C)crisprTF-4.1 與 PAM-NGT 之相對位置。(D)SpG 與 PAM-NGA 之相對位置。(E)crisprTF-4.1 與 PAM-NGA 之相對位置。

1. 1218 的胺基酸位置:並沒有改, originalTF、crisprTF-4.1 都是 G, 表示 1218 的位置 G 對 crisprTF-4.1 可否能認得到 **NGG** 十分重要
2. 1219 的位置由 originalTF 的 E 變成 C, 又與 SpG 的 Q, 性質相異, 可能是能辨認到 NAG 的原因
3. 由圖二十二(A)可知 originalTF-1219 的 E 與 PAM 序列 **NGG** 距離為 5.4 Å;圖(B)而 SpG-1219-Q 與 PAM 序列 **NGT** 距離為 4.1 Å, 與 PAM 的距離變近了;圖(C)crisprTF-4.1-1219-C 與 **NGT** 距離為 4.3 Å;圖(D)可知 SpG 的 1219-Q 與 PAM 序列 **NGA** 距離為 4.3 Å;圖(E)crisprTF-4.1-1219-C 與 **NGA** 距離為 4.3 Å;皆小於 5.4 Å, 表示有可能可以辨認到 **NGG**、**NGT**、**NGA**。

(二) xCas9

表七:xCas9、crisprTF-4.1、originalTF 之 1216-1221 胺基酸比較

	1216	1217	1218	1219	1220	1221
original crisprTF (NGG)	Ser	Ala	Gly	Glu	Leu	Gln
	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small	Polar Negative Charged Carboxylic	Non-polar Aliphatic Branched	Polar Nitrogen- containing Amides
xCas9 (NGN)	Ser	Ala	Gly	Val	Leu	Gln
	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Small Branched	Non-polar Aliphatic Branched	Polar Nitrogen- containing Amides
TF4.1 (NAG)	Trp	Cys	Gly	Cys	Ala	Gly
	Polar Aromatic Phenylic Nitrogen- containing	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small



圖二十三:(A)原始 crisprTF 與 PAM-NGG 之相對位置。(B)xCas9 與 PAM-NGA 之相對位置。(C)crisprTF-4.1 與 PAM-NGA 之相對位置。(D)xCas9 與 PAM-NGT 之相對位置。(E)crisprTF-4.1 與 PAM-NGT 之相對位置。

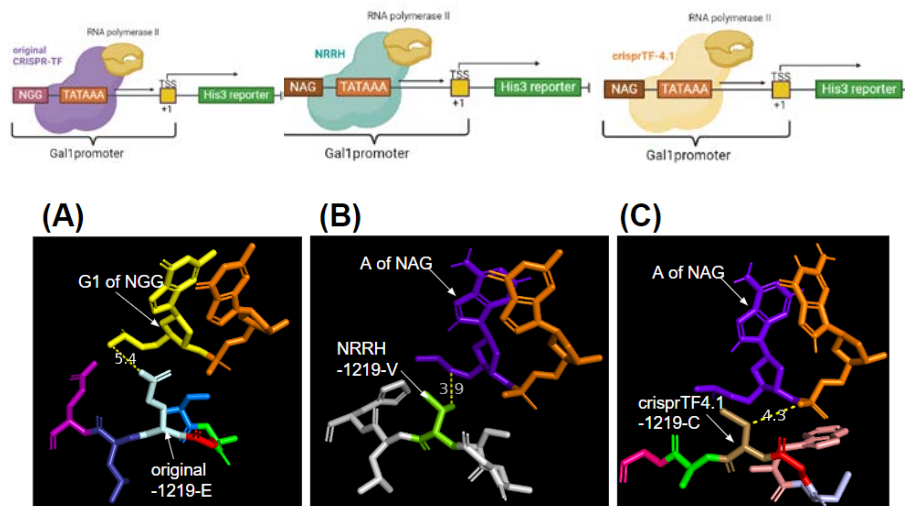
1. 1218 的胺基酸位置:並沒有改, originalTF、crisprTF-4.1 都是 G, 表示 1218 的位置 G 對 crisprTF-4.1 可否能認得到 NNG 的 G 十分重要
2. 由圖二十三(A)可知 originalTF-1219 的 E 與 PAM 序列 **NGG** 距離為 5.4Å;圖(B)而 xCas9-1219-Q 與 PAM 序列 **NGA** 離為 3.7Å、與 **NGA** 離為 3.8Å, 與 PAM 的距離變近了;圖(C)crisprTF-4.1-1219-C 與 **NGA** 距離為 4.1 Å; 圖(D)可知 xCas9-1219-V 與 PAM

序列 **NGT** 距離為 3.8 Å; 圖(E)crisprTF-4.1-1219-C 與 **NGT** 距離為 4.3 Å; 而 xCas9 可以辨認到 **NGN**，表示有可能可以辨認到 **NGG**、**NGA**、**NGT**。

(三) SpCas9-NRRH

表八:SpCas9-NRRH、crisprTF-4.1、originalTF 之 1216-1221 胺基酸比較

original crisprTF (NGG)	1216	1217	1218	1219	1220	1221
	Ser	Ala	Gly	Glu	Leu	Gln
	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small	Polar Negative Charged Carboxylic	Non-polar Aliphatic Branched	Polar Nitrogen- containing Amides
SpCas9- NRRH (NAC NAA NAG NGN)	Ser	Ala	Gly	Val	Leu	His
	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Small Branched	Non-polar Aliphatic Branched	Polar Positive Charged Aromatic Phenylic Nitrogen- containing
TF4.1 (NAG)	Trp	Cys	Gly	Cys	Ala	Gly
	Polar Aromatic Phenylic Nitrogen- containing	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small



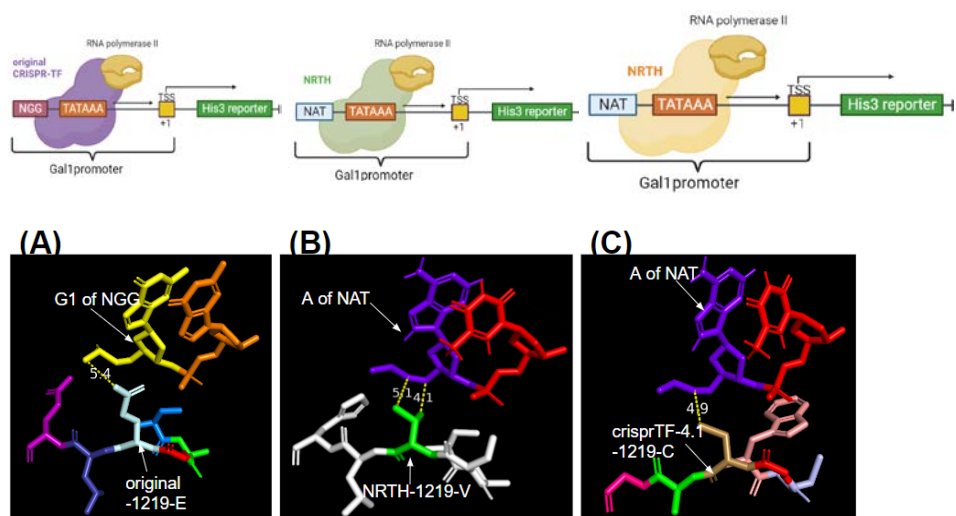
圖二十四:(A)原始 crisprTF 與 PAM-NGG 之相對位置。(B)SpCas9-NRRH 與 PAM-NAG 之相對位置。(C)crisprTF-4.1 與 PAM-NAG 之相對位置。

1. 1218 的胺基酸位置:並沒有改, originalTF、crisprTF-4.1 都是 G, 表示 1218 的位置 G 對 crisprTF-4.1 可否能認得到 NNG 的 G 十分重要
2. 由圖二十四(A)可知 originalTF-1219 的 E 與 PAM 序列 NGG 距離為 5.4 Å;圖(B)而 SpCas9-NRRH-1219-V 與 PAM 序列 NAG 離為 3.9 Å, 與 PAM 的距離變近了;圖(C)crisprTF-4.1-1219-C 與 NAG 距離為 4.3 Å; 與 PAM 的距離變近了, 而 SpCas9-NRRH 可以辨認到 NAC、NAA、NAG、NGN, 表示有可能可以辨認到 NGG、NAG。

(四) SpCas9-NRTH

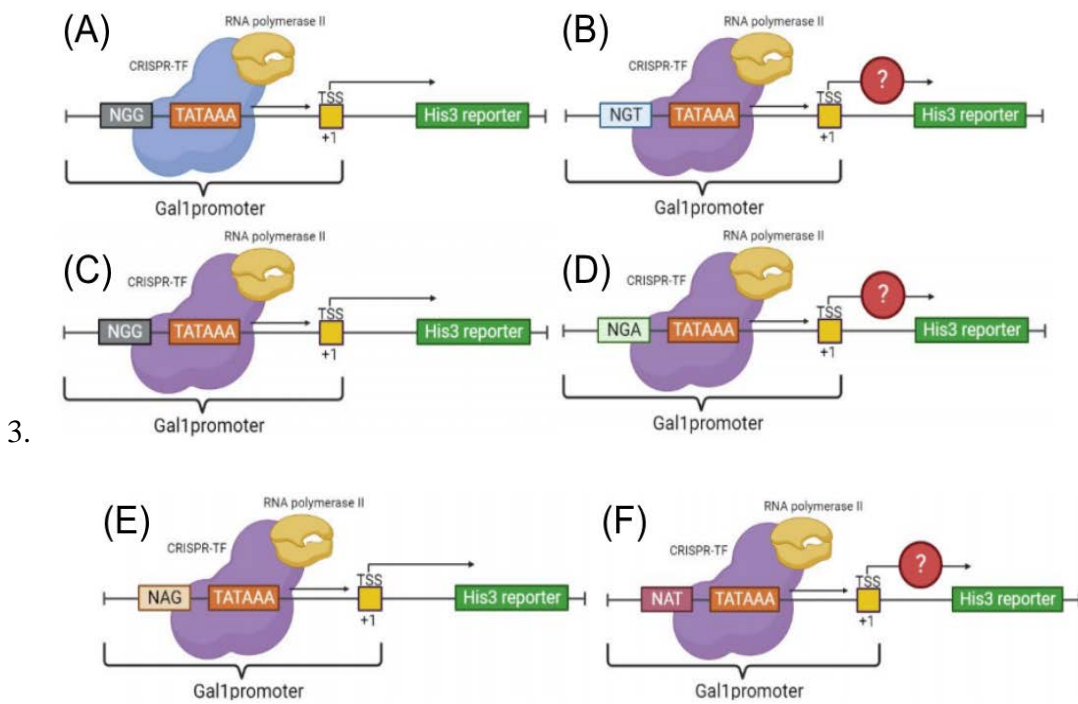
表九:SpCas9-NRTH、crisprTF-4.1、originalTF 之 1216-1221 胺基酸比較

original crisprTF (NGG)	1216	1217	1218	1219	1220	1221
	Ser	Ala	Gly	Glu	Leu	Gln
	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small	Polar Negative Charged Carboxylic	Non-polar Aliphatic Branched	Polar Nitrogen -containing Amides
SpCas9- NRTH (NAT NGN)	Ser	Ala	Ser	Val	Leu	His
	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Tiny Small	Polar Tiny Small Hydroxylic	Non-polar Aliphatic Branched Small	Non-polar Aliphatic Branched	Polar Positive Charged Aromatic Phenylic Nitrogen -containing
TF4.1 (NAG)	Trp	Cys	Gly	Cys	Ala	Gly
	Polar Aromatic Phenylic Nitrogen -containing	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Polar Tiny Small Sulphurous	Non-polar Aliphatic Tiny Small	Non-polar Aliphatic Tiny Small



圖二十五:(A)原始 crisprTF 與 PAM-NGG 之相對位置。(B)SpCas9-NRTH 與 PAM-NAT 之相對位置。(C)crisprTF-4.1 與 PAM-NAT 之相對位置。

1. 1218 的胺基酸位置:並沒有改變，originalTF、crisprTF-4.1 都是 G，表示 1218 的位置 G 對 crisprTF-4.1 可否能認得到 NGG 的 G 十分重要
2. 由圖二十五(A)可知 originalTF-1219 的 E 與 PAM 序列 NGG 距離為 5.4 Å;圖(B)而 SpCas9-NRTH-1219-V 與 PAM 序列 NAT 離為 4.1 Å，與 PAM 的距離變近了;圖(C)crisprTF-4.1-1219-C 與 NAT 距離為 4.9 Å;與 PAM 的距離皆變近了，而 SpCas9-NRTH 可以辨認到 NAT、NGN 表示有可能可以辨認到 NGG、NAT。



圖二十六:利用與不同 PAM 的結合驗證 crisprTF-4.1 的辨認能力。(A)(B)original crisprTF 和 crisprTF-4.1 辨認 NGG，能啟動 HIS3 基因。(C)實驗驗證後，crisprTF-4.1 能辨認 PAM-NAG 序列，啟動 HIS3 基因(D)(E)(F)經過文獻探討與討論，認為 crisprTF-4.1 能辨認 NGA、NGC、NAT。

伍、討論

一、Library

(一)Library DNA 長度:

Library DNA 事由 18 個隨機核苷酸組成，可轉譯成 6 個胺基酸

(二)文獻探討

我們發現研發出能打破 PAM 序列限制之研究中更改的 PI domain 胺基酸序列位置許多都集中在第 1219 和第 1333 個胺基酸附近，因此我們推測此二個位置可能對於控制 PI domain 辨認位點十分重要。例如:SpG 能辨認 PAM-NGN，其 PI domain 的序列改變為 **D1135L/S1136W/G1218K/E1219Q/R1335Q/T1337R**(Walton RT, et al., 2020)；抑或是 VRER 能辨認 PAM-NGCG，其 PI domain 的序列改變為 **D1135V/G1218R/R1335E/T1337R**。(Kleinstiver, B., et al., 2015)

(三)Library 庫的多樣性

LibraryDNA 為由 18 個隨機核苷酸組成的一段 DNA 片段，一共有 20 種胺基酸，因此最多有 20^6 種的 Library DNA。

(一)篩選出可能的 crisprTF v2

在建立 PI domain 資料庫後，我們進行 Library Yeast Transformation。由於此實驗要透過六個隨機胺基酸的組合(共有 18 個核苷酸序列，因此有 4^{18} 種組合)試圖得到一個能辨認為 NAG 的 PAM 序列的 PI domain，機率十分小，因此此實驗的效率就變得十分重要。**實驗能長出的總菌量越多，就越有機會得到我們想要的 crisprTF v2。**從(圖十四)能看出此次實驗菌落的總數約為 12,000 個菌落，代表我們在此次實驗中共從 4^{18} 種組合中分析了其中的 12000 種。

酵母菌轉型實驗中的 Linear DNA 是經 Afl2 和 Cla1 兩種限制酶剪切胺基酸序列第 1216 及第 1221 兩個切口的線狀 DNA。由於沒有完整接合且無 LibraryDNA 接合於 PI domain 的位置，無法辨認 PAM 序列，因此無論在何種培養基，生長數目皆極少；Linear DNA+Library DNA 是與六個隨機胺基酸序列在同源性重組後組合而成的質體 DNA。LibraryDNA 的隨機序

列中，若剛好有序列能使 PI domain 辨認 NAG 序列並啟動 CRISPR 系統，便能自行製造 Histidine，在 SCM-Ura -His 甚至是 SCM-Ura-His+Dox+3AT 培養基生長。經此實驗所得之酵母菌的 Linear DNA 和 Library DNA 有進行同源性重組，完整互補黏合。

(二)crisprTF v2 辨認的強度

在生長定量分析(Spotting assays)的(圖十六)中，在 Ura 營養缺陷的培養基中生長的菌落皆為擁有 Plasmid DNA 的酵母菌，因此得到的菌落數為此實驗的總菌落數。當酵母菌除了擁有 Plasmid DNA，其上的 Library DNA 序列還能使 PI domain 辨認為 NAG 的 PAM 序列，啟動 dCas9 的 CRISPR 系統，使其表現 HIS3 基因，我們就能在 SCM-Ura-His (缺乏 Ura 及 Histidine 營養的培養基)上篩選到此菌。(圖十六)中，紅框圈選起來的就是經處理後，目前看來能力最強，與 PAM-NAG 結合的最好的 crisprTF v2。

(三)確認獲得的細胞含有能辨認 PAM-NAG 序列的 crisprTF v2

1. 第二次生長定量分析(Spotting Assay)

藉由 PI domain 的定序，我們發現 crisprTF-3.1、crisprTF-3.2 以及 crisprTF4.2 皆為野生型的 crisprTF，我們並未成功更改他們的 PI domain，只有 crisprTF-4.1 有成功更改序列。(圖十八)雖然和正、負控制組相比，強度沒有正控制組強，但我們確實在(C)(D)(E)中看見菌落生長，顯示 crisprTF-4.1 具足夠強度且有能力藉由以新的 PI domain 和 PAM 序列-NAG 的結合，啟動下游基因。而結果看來最強的 4.3 經定序為 crisprTF-4.3，我們並未更改其序列。

2. 將 crisprTF v2 質體送入大腸桿菌(E.coli)並利用含有 AMP 培養基篩選

通過在有 Amp(Ampicillin)培養基上再次驗證其為 crisprTF v2 質體，因質體上帶有抗氨苄青黴素(AMP)基因(Li M, Liu Q, et al., 2019)，如果其質體為原本(圖十一)crisprTF v2 的質體，即我們的 candidate DNA，則轉殖進大腸桿菌之後，其細菌便可在加了氨苄青黴素(AMP)抗生素的培養基上存活。在過程中發現 crisprTF-1 無法生長，推測其轉殖入的質體並非 crisprTF v2，所以此 crisprTF-1 被排除，其他有生長的才得以進行後續實驗。

(圖十七)我們沒有在(A)組轉殖任何 DNA，大腸桿菌細胞沒有獲得任何質體。

因此，細胞無法在含有氨苄青黴素(AMP)的平板上存活。(B)組和(C)組中氨苄青黴

素(AMP)培養基上的大腸桿菌菌落具有其上具有 AMP 基因的質體，因此它們具有在有抗生素環境下存活的能力。我們使用這些菌落進行進一步的實驗，(D)組為本實驗的陽性對照。我們使用帶有 gRNA-NAG 和 AMP 基因的環狀質體。因此，所有帶有質粒的細胞都可以在含有氨苄青黴素(AMP)的培養基上存活。

5. 將大腸桿菌的質體抽出

crisprTF-2 在抽取質體時失敗，而 crisprTF-5 的質體大小不符，所以這些 candidate 無法在後續實驗中使用，因此剩 crisprTF-3、crisprTF-4 是合格的。

三、從 PI domain 被更改的胺基酸序列分析其功能與改變

(一)定序結果

由(圖十二)可知，我們成功將 crisprTF-4.1 的 PI domain 第 1216-1221 個胺基酸序列由 **SAGELQ** 更換為 **WCGCAG**，六個胺基酸中有五個改變為不同的胺基酸序列。第 1218 個胺基酸和原始的 crisprTF 相同，仍為 G(Glycine)。因此，我們推測這個胺基酸應該對於 PI domain 的辨認 NGG 的第二個 G 及 NAG 的 PAM 序列中的 G 十分重要且不可或缺，才會即使更換序列仍換成相同的胺基酸。

此外，我們也推測其他五個更換的胺基酸可能對於 PAM 序列的第二個核苷酸的辨認(也就是 NAG 序列的 A 以及 NGG 序列的第一個 G)十分重要。觀察改變的胺基酸序列，我們發現第 1216 個胺基酸由 S(Serine)更換為 W(Tryptophan)，少了一個氫鍵且較大，第 1217 及 1219 個胺基酸分別從 A(Alanine)和 E(Glutamate)更換為 C(Cystine)，都較比原始多獲得了一個雙硫鍵，這些結構的改變都可能是影響了 PI domain 和 PAM 序列結合的原因。

透過分析這些結構與序列，我們推測或許未來能只改變其中五個甚至更少的胺基酸序列來改變 PI domain 的辨認位置。

四、多樣化發展的 CRISPR

由於 CRISPR 擁有能利用 gRNA 與 PAM 序列結合辨認切點並準確剪切的特性，因此自此系統被發現以來，CRISPR/Cas9 漸漸成為在基因編輯廣泛被使用的技術，使我們能準確且有效編輯細胞內的部分基因。

現今，CRISPR 技術已成為針對目標基因功能專一的研究工具，我們甚至能直接向公司訂購更改序列後的質體，為學術研究增添許多便利性。此外，此技術也廣泛使用於馬鈴薯、蘑菇、玉米等農產品的改良、HIV 疾病治療的研究(Xun, J.,et al.,2021)，甚至是癌症的治療(Katti, A.,et al.,2021)。

目前我們已成功破除 PAM 序列(NGG)的限制，實驗證實 crisprTF-4.1 能辨認 PAM-NAG 序列，經過文獻探討指出 crisprTF-4.1 亦可能辨認 NGA、NGC、NAT 的 PAM 序列。未來若能破除更多序列的限制，必定能對此技術的發展帶來便利與貢獻。

陸、結論與應用

一、結論

(一)在 CRISPR 系統中 PI domain 的序列 1216-1221 對於辨認 PAM 序列的功能具有重要性。

(二)過去的 CRISPR 之所以無法成功，例如:找出修復 DNA 的基因組，是因 PAM-NGG 的限制。透過我們的實驗研發 crisprTF v2，能將 crisprTF-NAG 用於將來 CRISPR 的應用。

(三)我們成功的編輯 PI domain 中的序列 1216-1221，且藉由定序得知了序列 1216-1221 中被我們由(SAGELQ)修改成(WCGCAG)

二、持續進行的實驗

(一)利用酵母菌驗證 crisprTF-4.1 能辨認 NGG、NAG 以外的其他 PAM 序列。

(二)將 crisprTF-4.1 應用於有 MMS 損傷的酵母菌細胞中，希望辨認能力更強更廣的 crisprTF v2 能修復細胞。

三、未來展望

我們在研究中已使 crisprTF v2 成功辨認 PAM 序列-NAG，未來我們除了要將研發的 crisprTF v2 做實際的應用，更希望能利用相同的研究方式與技術研發不同版本的 crisprTF v2 使其有能力辨認 NGG、NAG 以外其他 14 種 PAM 序列。

PAM		NNG			
		A	T	C	G
NGN	A	AA	AT	AC	AG
	T	TA	TT	TC	TG
	C	CA	CT	CC	CG
	G	GA	GT	GC	GG

圖二十七:橘底的方格為我們設計的 crisprTF v2，可以與 NAG 的 PAM 序列結合。紅底方格是原始的 PRISM v1，能與 NAG 的 PAM 序列結合，其他 14 個序列能以此研究的相同方法破解。

柒、參考文獻資料

- 1、 Adli, M, “The CRISPR tool kit for genome editing and beyond,” Nat Commun 9, 1911 (May 15, 2018).
<https://doi.org/10.1038/s41467-018-04252-2>
- 2、 Chen YC, Farzadfard F, Gharaei N, Chen WCW, Cao J, Lu TK, “Randomized CRISPR-Cas Transcriptional Perturbation Screening Reveals Protective Genes against Alpha-Synuclein Toxicity,” Volume 68, issue 1, P247-257.E5, (October 5, 2017) . <https://doi.org/10.1016/j.molcel.2017.09.014>
- 3、 Wang Y, Jensen RC, Stumph WE. “Role of TATA box sequence and orientation in determining RNA polymerase II/III transcription specificity. “Nucleic Acids Res. (August 1, 1996) <https://doi.org/10.1093%2Fnar%2F24.15.3100>
- 4、 Xun, J., Zhang, X., Guo, S. et al. “Editing out HIV: application of gene editing technology to achieve functional cure,” Retrovirology 18, 39 (December 18, 2021). <https://doi.org/10.1186/s12977-021-00581-1>
- 5、 Katti, A., Diaz, B.J., Caragine, C.M. et al. “CRISPR in cancer biology and therapy,” Nat Rev Cancer 22, 259–279 (February 22, 2022). <https://doi.org/10.1038/s41568-022-00441-w>
- 6、 Gasiunas G , Barrangou R , Horvath P , Siksnys V. et al. “Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.” Proc Natl Acad Sci U S A. (September 4, 2012).
<https://doi.org/10.1073/pnas.1208507109>
- 7、 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E, “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity,” Science, Vol 337, Issue 6096 (September 4, 2012).
<https://www.science.org/doi/10.1126/science.1225829>
- 8、 Park JJ, Dempewolf E, Zhang W, Wang ZY, “RNA-guided transcriptional activation via CRISPR/dCas9 mimics overexpression phenotypes in Arabidopsis,” PLOS ONE 12(6): e0179410. (June 16, 2017).
<https://doi.org/10.1371/journal.pone.0179410>
- 9、 Gleditsch D, Pausch P, Müller-Esparza H, Özcan A, Guo X, Bange G, Randau L, “PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures,” RNA Biology, 16:4, 504-517, (September 18, 2018).
<https://doi.org/10.1080/15476286.2018.1504546>
- 10、 Mitsuyoshi Ueda, Hiroshi Kinoshita, Tomoko Yoshida, Naomi Kamasawa, Masako Osumi, Atsuo Tanaka, “Effect of catalase-specific inhibitor 3-amino-1,2,4-triazole on yeast peroxisomal catalase in vivo,” FEMS Microbiology Letters, Volume 219, Issue 1, Pages 93–98, (February 1, 2003). [https://doi.org/10.1016/S0378-1097\(02\)01201-6](https://doi.org/10.1016/S0378-1097(02)01201-6)
- 11、 Daan J.W. Brocken, Mariliis Tark-Dame and Remus T. Dame, “dCas9: A Versatile Tool for Epigenome Editing” Curr. Issues Mol. Biol. 26: 15-32. (September 7, 2017). <https://doi.org/10.21775/cimb.026.015>
- 12、 Taymaz-Nikerel H, Karabekmez ME, Eraslan S, Kırdar B, “Doxorubicin induces an extensive transcriptional and metabolic rewiring in yeast cells,” Sci Rep 8, 13672. (September 12, 2018)<https://doi.org/10.1038/s41598-018-31939-9>
- 13、 Li M, Liu Q, Teng Y, Ou L, Xi Y, Chen S, Duan G, “The resistance mechanism of Escherichia coli induced by ampicillin in laboratory. Infect Drug Resist,” (September 11, 2019) <https://doi.org/10.2147/idr.s221212>
- 14、 Miller, S.M., Wang, T., Randolph, P.B. et al. “Continuous evolution of SpCas9 variants compatible with non-G PAMs. “ Nat Biotechnol 38, 471–481 (February 10,2020)<https://doi.org/10.1038/s41587-020-0412-8>
- 15、 Hu JH, et.al. “Evolved Cas9 variants with broad PAM compatibility and high DNA specificity.” Nature. ;57-63.(April 5,2018)<https://doi.org/10.1038/nature26155>
- 16、 Liu, Z., et al. “Expanded targeting scope and enhanced base editing efficiency in rabbit using optimized xCas9(3.7).” Cell. Mol. Life Sci. 76, 4155–4164 (April 27, 2019). <https://doi.org/10.1007/s00018-019-03110-8>

- 17、 Kleinstiver, B., et al. "Engineered CRISPR-Cas9 nucleases with altered PAM specificities." *Nature* 523, 481–485 (June 22, 2015). <https://doi.org/10.1038/nature14592>
- 18、 Walton RT, Christie KA, Whittaker MN, Kleinstiver BP. "Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants." (March 26, 2020). <https://doi.org/10.1126/science.aba8853>

【評語】 052012

CRISPR 是相當新的基因編輯技術，作者們描述了更改後的 PAM 相互作用區序列以及可能辨認的其他 PAM 序列，希望讓這個技術變得更好用。原先用意很好，研究也尚未完成，很值得持續探討。作者對於篩選的原理，可以做更詳盡的說明。目前用粗略的細胞生長來代表蛋白與 DNA 的結合強度，做初步篩選可以，但是並不代表實際可行。仍需有進一步的實驗證據來支持。

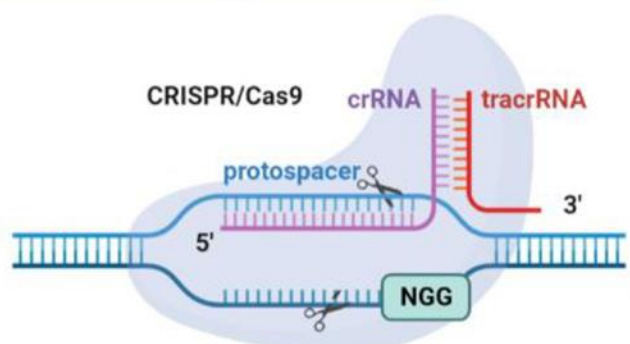
作品海報

簡介

CRISPR/Cas9是合成生物學中的重要技術，被廣泛使用於基礎研究、基因調控、基因編輯、藥物研發與醫學治療方法等。在CRISPR/Cas9技術中，需引導RNA(guide RNA)將Cas9核酸酶帶至目標DNA序列後，Cas9核酸酶會辨認特定的Protospacer adjacent motif(PAM)序列-NGG (N為A、T、C、G任意核苷酸；G為鳥嘌呤)，切割緊鄰的雙股DNA。然而，這個特定的PAM核酸序列(NGG)卻會因此限制了其可編輯的基因位點。為了使CRISPR技術的應用更廣泛且便利，我們研發出新一代的crisprTF v2，以NAG為起點突破PAM核酸序列(PAM Sequence)對於特定序列(NGG)的要求，並透過定序得知更改後的PAM 相互作用區域(PAM Interaction domain,簡稱為PI domain)序列為何，經驗證為PAM-flexible variant，能為基因工程的研究帶來便利性與更多的可能。

研究背景

原始的crisprTF v1



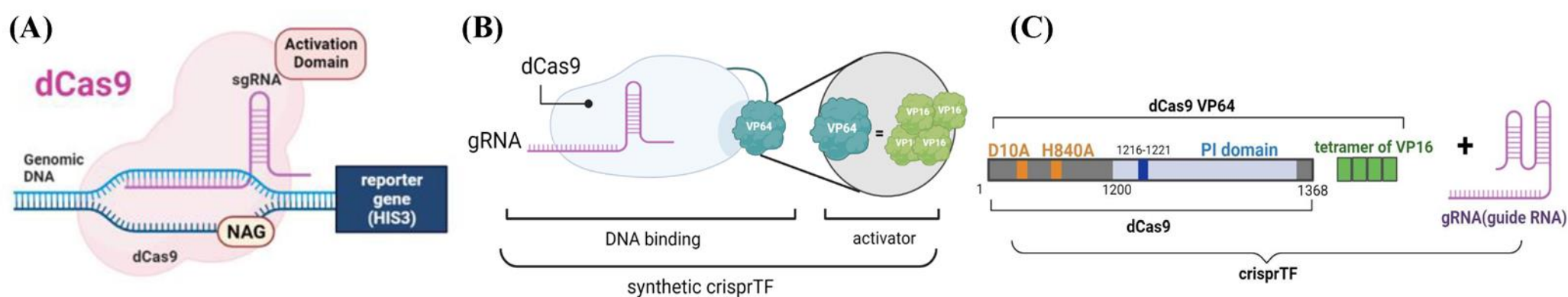
圖一:crisprTF-WT-NGG示意圖。[1]

- 在原始 CRISPR/Cas9 中，tracrRNA 與 crRNA 組成的 gRNA 會辨認 protospacer，也就是20個序列，而PI domain則會辨認目標基因下游的 PAM 序列(NGG)，其中，N 為 ATCG 任一核苷酸。接著，便會啟動 CRISPR 機制，利用核酸內切酶剪切目標基因序列。[1]
- PAM 序列(Protospacer Adjacent Motif Sequence) 是一段2至5個鹼基對的序列，位於目標DNA旁。

實驗設計

建構dCas9-crisprTF

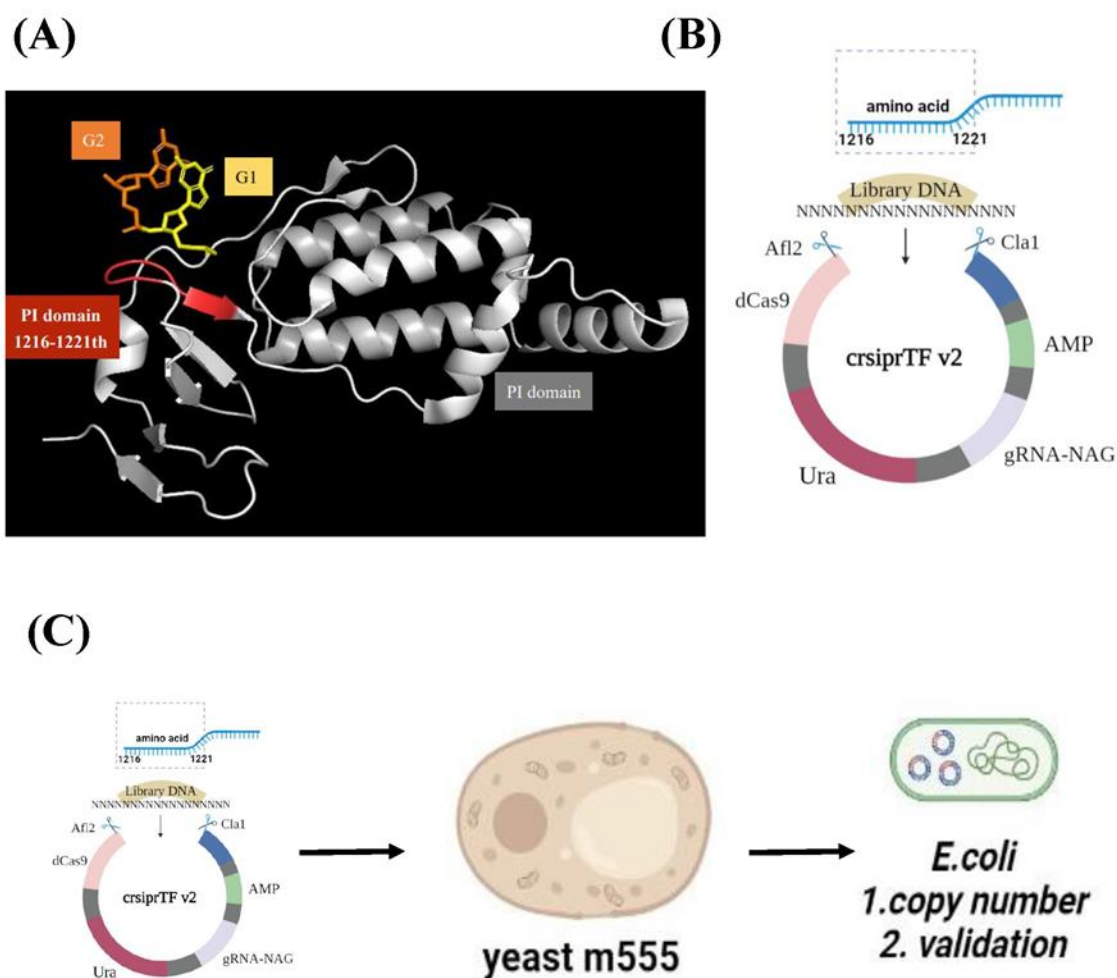
CrisprTF(crispr Transcription Factor)含有dCas9 (deadCas9),VP64和轉錄因子，其中dCas9的核酸內切酶失去活性，因此失去剪切能力。[2]



圖二:人工合成的dCas9。(A) CRISPR/dCas9機制示意圖。當dCas9識別正確的序列時，標記為藍色的reporter基因會表現HIS3基因。因此，細胞可以自行製造Histidine，並在缺少Histidine的培養基上生長。(B) crisprTF之組成。VP64是由4個VP16所組成，VP16為轉錄因子，可啟動基因表現的調控。(C) crisprTF的序列組成。Cas9 蛋白有兩個 DNA 核酸內切酶結構域，HNH (或 McrA) 和 RuvC 結構域，每個結構域都賦予對目標雙鏈 DNA 的一條鏈的切割活性。這些結構域中催化殘基的胺基酸取代，HNH 結構域的 D10A 和 RuvC 樣結構域的H840A，消除了 DNA 切割活性，同時保留了其特定的 DNA 結合活性。[3]

實驗結果與討論

以N序列產生Library DNA



- 本研究利用CRISPR/dCas9改良出一套能夠辨認到NGG以外序列(如:NAG序列)的系統，篩選出能夠被gRNA(NAG)辨認，開啟下游基因的序列，開創crisprTF v2。
- 將帶有AMP、gRNA及Ura等基因且被Afl2及Cla1兩種限制酶切開的環狀DNA以及Library DNA(即dCas9辨認PAM序列之PI-domain上第1216-1221序列更改為隨機的六個胺基酸序列組合)希望將質體與Library DNA送入酵母菌細胞內後，能得到可使PI-domain辨認NAG序列的能力，並啟動Histidine基因表現的轉殖酵母菌。

圖三: PI domain Library的建構及轉殖。

(A)透過PyMOL模擬分析發現PI domain的1216-1221序列離PAM最近可能是辨認PAM重要位置。(B)crisprTF的質體，此質體帶有AMP、URA3等基因以及dCas9和gRNA(NAG)。黃色為隨機核苷酸序列(Library DNA)。(C)本實驗DNA轉殖及檢驗流程。

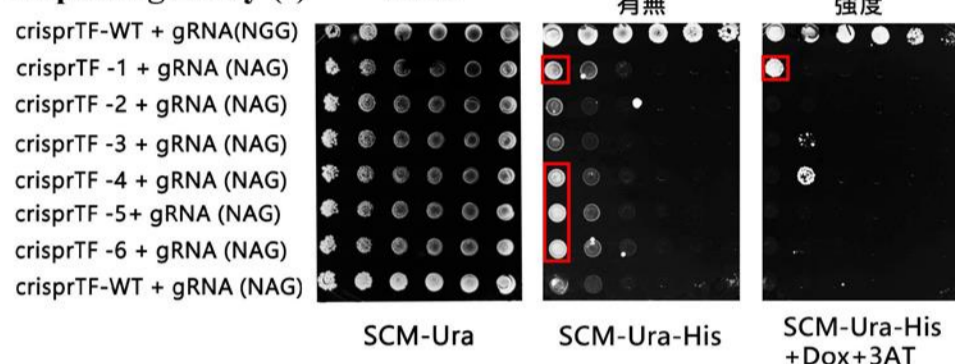
從PI domain資料庫篩選crisprTF v2

成功建立PI domain資料庫(PI domain Library)後，將隨機序列組合與載體接合，篩選出含有能使PI domain辨認到PAM-NAG之重組DNA的酵母菌。

- SCM(Synthetic complete mix)含有酵母菌生長的基本營養，但不含Uracil, Leucine, Tryptophan等胺基酸。
- -U表示培養基未加入Uracil，-H表示未加入Histidine，+Dox為添加 doxycycline，+3AT為添加 3-amino-1,2,4-triazole。
- 表一(A)可得實驗總菌落數，菌落小，但密集且多，因此較有機會從中得到可以辨認PAM-NAG的酵母菌。
- 表一(B)(C)(D)皆在篩選crisprTF v2。其中(D)為擁有修改過的質體，既能辨認NAG序列，啟動CRISPR系統，又有足夠生存能力的菌落。

尋找PAM結合能力強的crisprTF v2

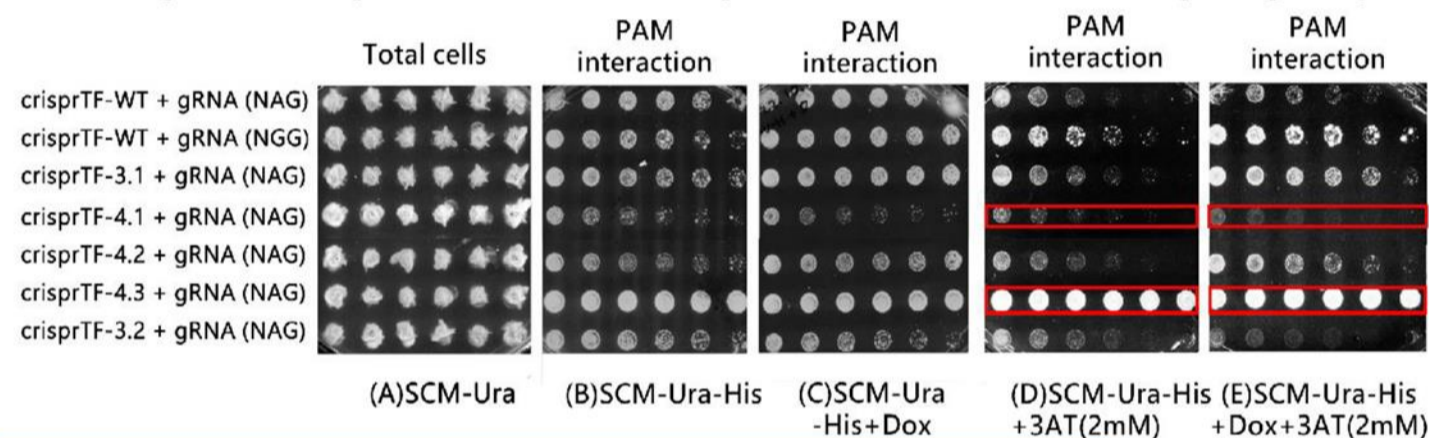
1. Spotting Assay (I)



2. crisprTF-3、crisprTF-4 Spotting Assay (II)

將DNA轉殖入E.coli確認質體為crisprTF v2並大量複製DNA、確認crisprTF v2的DNA大小後，再次將DNA轉殖入酵母菌確認。

➢ 認為crisprTF-3及crisprTF-4為最有可能的crisprTF v2，因此我們再次進行Spotting Assay驗證，並抽DNA定序確定。



圖五:crisprTF-3、crisprTF-4辨認PAM-NAG的強度分析。

圖中可以从生長情況看出辨認的強度差異，crisprTF-4.1確認有改到序列，為crisprTF v2；crisprTF-3、crisprTF-4.2、crisprTF-4.3經定序，確定未更改到序列，為原始的crisprTF-WT-NGG。

PI domain的第1218個胺基酸Glycine對辨認NAG的G有重要地位

➢ 分析crisprTF-4.1之胺基酸序列，我們發現即使序列更換，第1218個胺基酸序列仍對應到G(Glycine)，可推知其可能對辨認NAG的G和NGG的第二個G有重要地位。

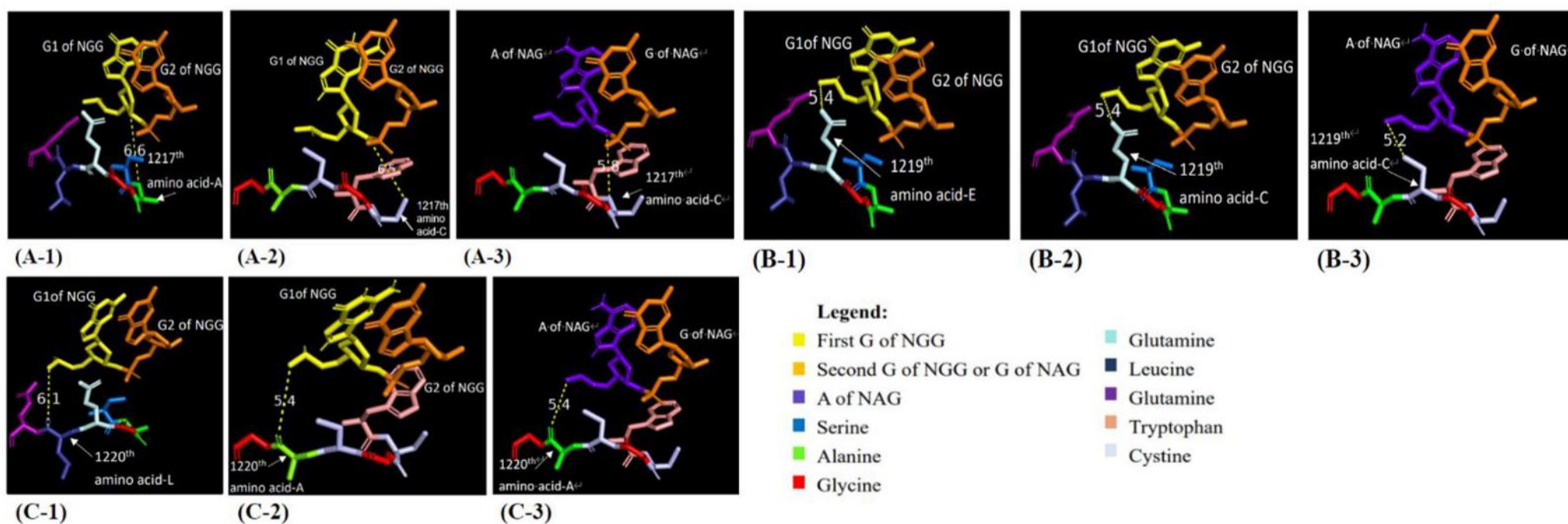


圖六:從軟體SnapGene Viewer擷取的序列圖。crisprTF-WT-NGG之PI domain第1216-1221胺基酸序列為SAGELQ，crisprTF-4.1更改為WCGCAG。

測量PI domain與PAM序列的距離

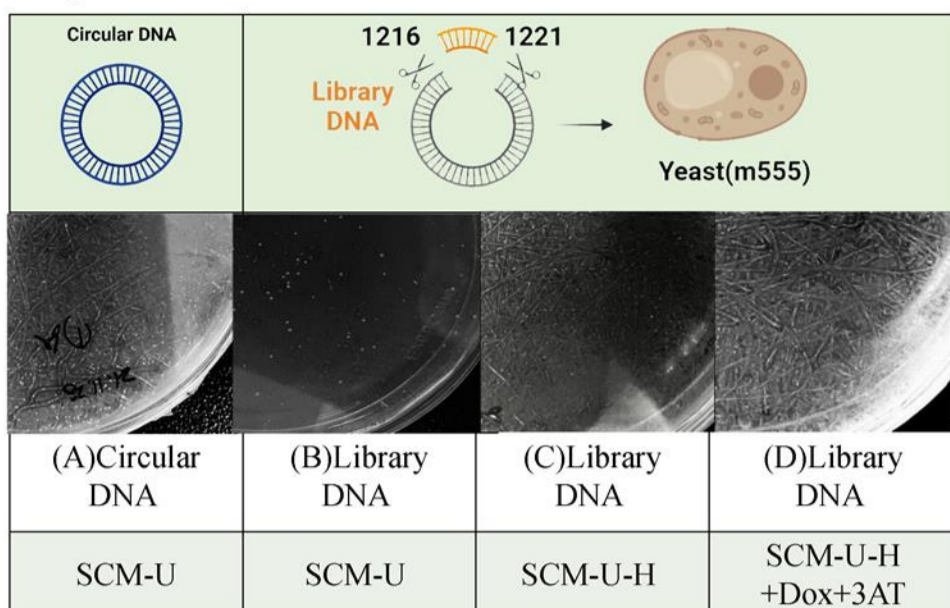
➢ 以PI domain第1217(A)、1219(B)與1220(C)個胺基酸為例，模擬後發現PI domain與PAM序列的距離皆變短了:分別從6.6Å改為5.8Å、5.4Å改為5.2Å、6.1Å改為5.4Å

➢ PI domain與PAM序列的距離變近可能是使crisprTF-4.1能辨認PAM-NAG可能的原因。



圖七:利用PyMOL模擬crisprTF與PAM序列之互動，並測量兩者間的距離。(A-1)(B-1)(C-1)是crisprTF-WT-NGG；(A-2)(B-2)(C-2)是crisprTF-4.1-NGG；(A-3)(B-3)(C-3)是crisprTF-4.1-NAG。

表一:Yeast Library Transformation實驗結果。將隨機序列與載體送入酵母菌，使其進行同源性重組，並利用培養基篩選具辨認PAM-NAG的菌落。(B)(C)(D)生存的菌落為可能的crisprTF v2。



從表一(D)挑選6個菌落crisprTF-1~6進行強度分析之結果，能確定其正確性及辨認能力。

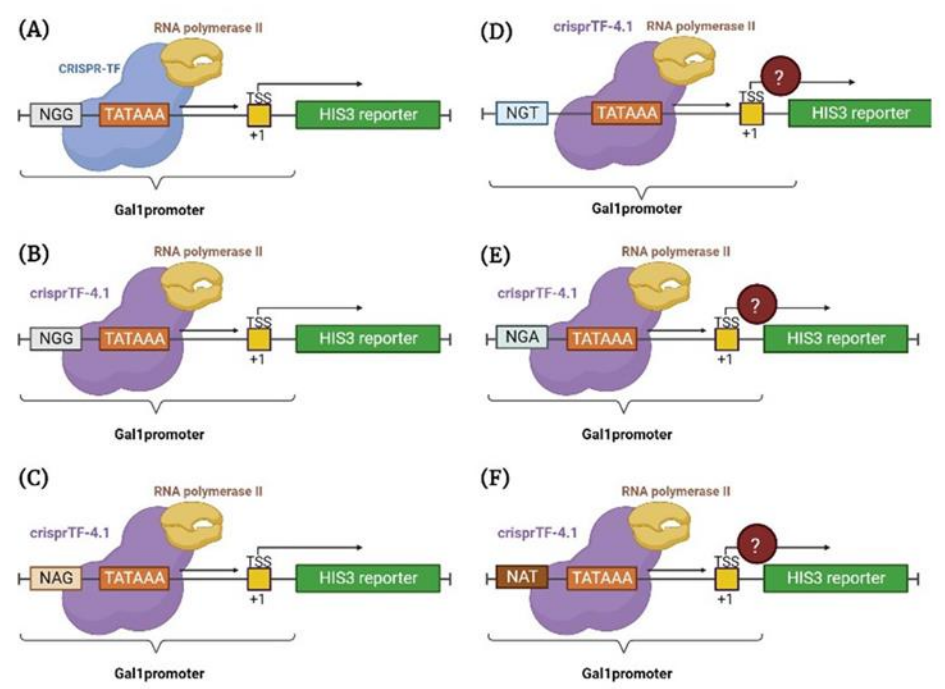
圖四:分析crisprTF v2辨認PAM序列的強度差異。

透過文獻探討crisprTF4.1能辨認其他PAM-flexible variant序列的可能性

表二:crisprTF-WT-NGG與文獻中PAM variants、crisprTF-4.1之胺基酸序列比較以及其能辨認的PAM序列。[4]

	1216-1221 (紅字表示更改到的胺基酸序列)	PAM序列
crisprTF-WT-NGG	SAGELQ	NGG
SpG	SAKQLQ	NGN
xCas9	SAGVLQ	NGN
*SpCas9-NRRH	SAGVLH	NAC NAA NAG NGN
SpCas9-NRTH	SASVLH	NAT NGN
crisprTF-4.1	WCGCAG	NGG NAG

*N:A,T,C,G任意核苷酸 *R:A or G *H:A or C or T

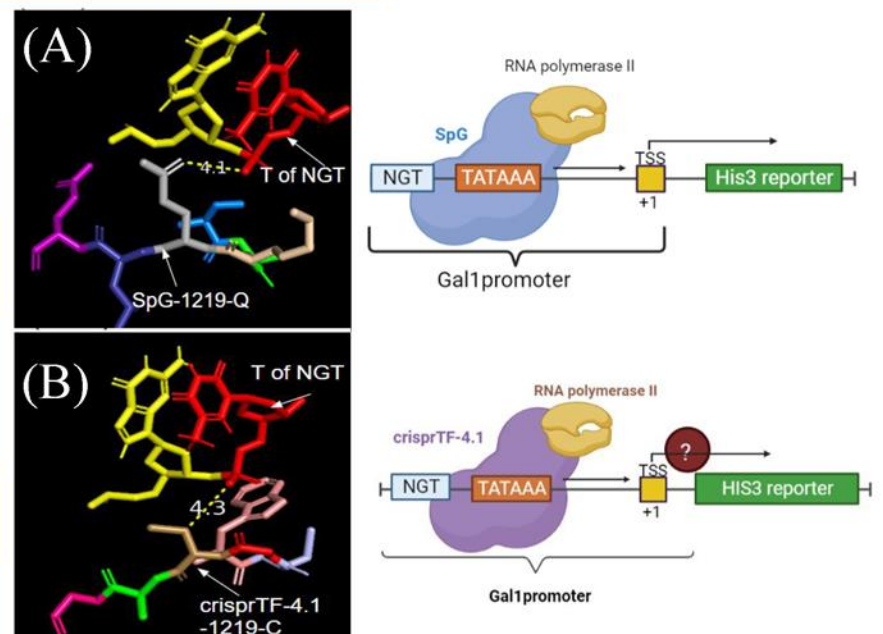


圖八:crisprTF-4.1與可能結合的PAM。(A)crisprTF-WT+PAM(NGG) (B)crisprTF-4.1+PAM(NGG) (C)crisprTF-4.1+PAM(NAG) (D) crisprTF-4.1+PAM(NGT) (E)crisprTF-4.1+PAM(NGA) (F)crisprTF-4.1+PAM(NAT)

以胺基酸1219與PAM-NGT為例討論crisprTF-4.1能辨認其他PAM的可能性

表三:crisprTF-WT-NGG與文獻中PAM variants、crisprTF-4.1之胺基酸特性及DNA距離比較。

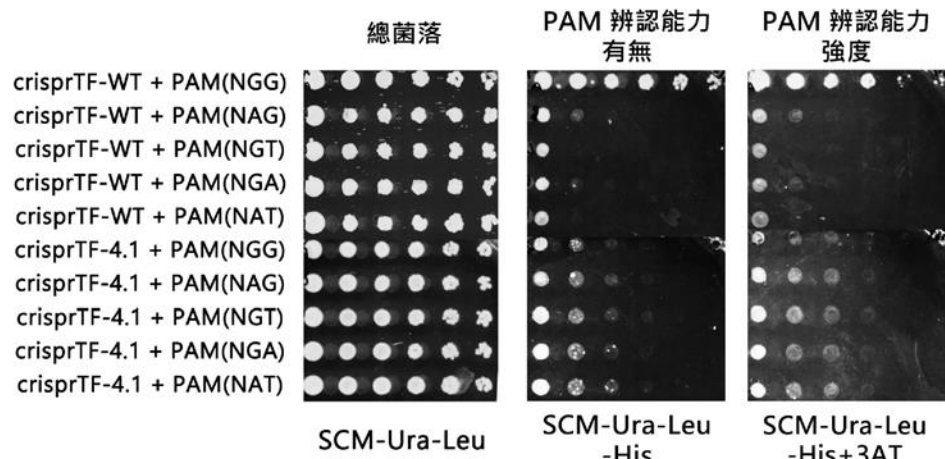
	crisprTF-WT	crisprTF-4.1	SpG	xcas9	SpCas9-NRRH
1219	Glutamic acid	Cysteine	Glutamine	Valine	Valine
胺基酸特性	Polar Negative Charged Carboxylic	Polar Tiny Small Sulphurous	Polar Nitrogen-containing Amides	Non-polar Aliphatic Small Branched	Non-polar Aliphatic Small Branched
PAM序列	NGG	△待驗證	NGN	NGN	NAC NAA NAG NGN
與PAM-NGT距離	5.4Å	4.3Å	4.1Å	3.8Å	-



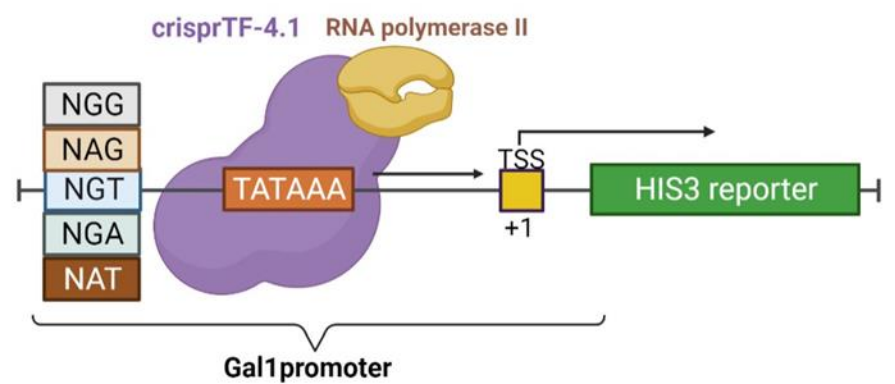
圖九:胺基酸序列第1219與NGT的距離比較。(A) SpG-NGT(B)crisprTF-4.1-1219-NGT。

以Spotting assay驗證crisprTF-4.1辨認PAM-flexible variant的強度

進行Spotting Assay的菌落由SCM-Ura-Leu培養基上的挑選，實驗設置與過去的Spotting Assay相似。A為Positive control，B~E為Negative control，因為crisprTF-WT-NGG只辨認PAM(NGG)。F~J為Candidate，生長能力較Negative control強，且即使加入1mM 3AT仍能辨認並生長。



圖十: crisprTF-4.1+gRNA(NAG)辨認五種PAM的強度分析。



圖十一: crisprTF-4.1能辨認五種PAM序列NGG、NAG、NGT、NGA、NAT，為PAM-flexible variant。

結論

1. 過去的CRISPR之所以受到限制是因PAM-NGG的限制。透過本研究研發crisprTF v2，能擴大CRISPR的應用。
2. 此方法突破CRISPR對於PAM-NGG序列的限制，使得新的PI domain能具有PAM-NAG序列的辨認能力。
3. 藉由定序得知了PI domain胺基酸序列第1216-1221中由SAGELQ修改成WCGCAG。
4. 更改後的crisprTF v2與PAM序列的距離變小，可增加兩者交互作用的機率。
5. 此研究成果顯示crisprTF v2可辨認到五種PAM(NGG、NAG、NGT、NGA、NAT)

未來展望與應用

➤ 持續進行的實驗

- (一) ChIP PCR(Chromatin Immunoprecipitation Polymerase Chain Reaction) 使用抗體確認結合的PAM序列，並找出調控基因。
- (二) 利用crisprTF-4.1辨認其他11種PAM序列。
- (三) 應用crisprTF v2在受到DNA損傷的酵母菌細胞。

若得到生存細胞，表示crisprTF v2有機會運用在基因損傷疾病的治療上。

➤ 未來展望:

1. 此研究中已使crisprTF v2成功辨認PAM序列-NGG、NAG、NGT、NGA、NAT，未來我們希望能利用相同的研究方式與技術研發不同版本的PRISM使其有能力辨認NGG、NAG、NGT、NGA、NAT以外11種PAM序列，完全打破此技術的限制。
2. 在DNA損傷的修復系統中加入CRISPR，讓細胞在受損害時，即使調控網絡失常，無法發揮作用，細胞仍能自行修復並且找出參與DNA損傷修復機制的基因組。

參考資料與文獻

1. Adli, M. (2018), The CRISPR tool kit for genome editing and beyond. Nat Commun 9.
2. Daan J.W. Brocken, Mariliis Tark-Dame and Remus T. Dame, (2018), dCas9: A Versatile Tool for Epigenome Editing.
3. Park JJ, Dempewolf E, Zhang W, Wang ZY, (2017)RNA-guided transcriptional activation via CRISPR/dCas9 mimics over expression phenotypes in Arabidopsis.
4. Miller, S.M., Wang, T., Randolph, P.B. et al, (2020). Continuous evolution of SpCas9 variants compatible with non-G PAMs. Nat Biotechnol 38.