

有效計算智慧為主的方法用於誘發突變引子設計

鄭煜輝
朝陽科技大學
資訊與通訊系副教授
yuhuei.cheng@gmail.com

郭哲男
稻江科技暨管理學院
動畫遊戲設計學系助理教授
fkikimo@hotmail.com

摘要

在本研究中，提出一種有效計算智慧為主的方法來設計聚合酶鏈鎖反應-限制性片段長度多型性 (polymerase chain reaction-restriction fragment length polymorphism, PCR-RFLP) 的誘發突變引子，用於單核苷酸多型性 (single nucleotide polymorphism, SNP) 基因定型。一種具效率且基於 REBASE 之限制酶搜尋的新方法被引入到該方法中。誘發突變矩陣被用於提升獲得目標單一核苷酸多型性限制酶的效率。此外，較準確的熱力學解鏈溫度計算也被應用於該方法。我們已經利用所提出的方法設計於具有錯配 PCR-RFLP 的 25 個目標單一核苷酸多型性上。結果顯示此方法有助於設計符合條件的誘發突變引子與獲得目標單一核苷酸多型性上可用的限制酶。

關鍵詞：計算智慧、聚合酶鏈鎖反應-限制性片段長度多型性、誘發突變引子、單一核苷酸多型性、限制酶。

Abstract

In this study, we propose an efficient computational intelligence-based approach to design mutagenic primers for single nucleotide polymorphism (SNP) genotyping in polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP). A new method of restriction enzyme search based on REBASE is introduced into the method. A mutagenic matrix is used to increase the efficiency of obtaining restriction enzymes for the target single-nucleotide polymorphism. In addition, an accurate thermodynamic melting temperature calculation is also applied to the method. There are 25 target single nucleotide polymorphisms with mismatch PCR-RFLP

have been performed by the method. The results show that the proposed method facilitates the design of mutagenic primers that satisfied the desired conditions and get available restriction enzymes for the target SNP.

Keywords: Computational intelligence, PCR-RFLP, mutagenic primer, SNP, restriction enzyme.

1. 前言

單一核苷酸多型性 (single nucleotide polymorphism, SNP) 是最廣泛的遺傳變異，並且是重要的遺傳標記。聚合酶鏈鎖反應-限制性片段長度多型性 (polymerase chain reaction-restriction fragment length polymorphism, PCR-RFLP) 是一種快速和經濟的生物技術，經常用於 SNP 基因定型上[1]。然而，許多 SNPs 因為缺乏可用的限制酶，而無法用於 PCR-RFLP 實驗進行基因定型，我們稱此為錯配 PCR-RFLP。幸運的是，誘發突變引子設計可以解決這個問題[2]。

在相關研究中，Cheng 等人在 2013 年使用具有固定和自適應停止標準的遺傳算法 (GA) 設計 PCR-RFLP 引子[3]。此外，他們還在 PCR-RFLP 引子設計中評估 GA 的適當參數[4]。2016 年，Cheng 等人提出了一種改進 GA 的新方法[5]。然而，這些方法不適合於錯配 PCR-RFLP。有數個系統提供誘發突變引子和限制性酶的資訊，包括 SNP Cutter [6]，V-MitoSNP [7] 和 Prim-SNPing [8]。但是，SNP Cutter 僅限於較舊的資料庫，許多較新的 SNP ID 無法使用，並且僅有幾種限制酶可供選擇利用。V-MitoSNP 只提供人類線粒體 SNP (mtSNPs) 的資訊。Prim-SNPing 受其方法策略的限制，而大多提供無效的引子。因此，誘發突變引子和可用的限制酶難以獲得，而使得 PCR-RFLP 實驗無法進行。

2012 年，基於 GA 的方法首次被提出用於錯配 PCR-RFLP 的誘發突變引子設計上[2]。解鏈溫度利用 $2 \times \text{AT} + 4 \times \text{CG}$ 規則計算，即 Wallace 的公式[9]，用於設計誘發突變引子。然而，即使許多文獻使用 Wallace 的公式設計引子[10-12]，但是許多研究人員認為它仍然為一種不準確的方法。SantaLucia [13]提出的熱力學 NN (Nearest Neighbor)模型被認為是用於解鏈溫度估計的精確方法。有許多軟體和工具使用 SantaLucia 的解鏈溫度計算公式，如 Primer3 [14]、MethPrimer [15]、BatchPrimer3 [16]、QuantPrime [17]和 Primer-BLAST [18]。

在本研究中，我們提出了一種有效計算智慧的方法，應用學習的概念[19-21]設計誘發突變引子。透過改變目標 SNP 鄰近的核苷酸[22, 23]並基於 REBASE 限制酶搜尋的新方法來發現可用的限制酶。本研究方法提供了符合經驗之 PCR-RFLP 實驗的多項引子限制，包括引子長度、引子長度差、引子解鏈溫度、引子解鏈溫度差、GC 比例、GC 鉗夾、交叉二聚體、自身二聚體、髮夾結構、特異性和 PCR 產物大小[12, 24-26]。本研究已經使用 25 個具有錯配 PCR-RFLP 的目標 SNPs 進行測試。結果顯示所提出的方法是一種有用的方法，可為目標 SNP 提供可用的限制酶和設計符合條件的誘發突變引子，以應用於 PCR-RFLP 實驗上。

2. 方法

2.1 誘發突變引子設計的定義

誘發突變引子設計是藉由改變序列中的核苷酸，找出可用的限制酶以辨認序列中的目標 SNP，同時也為 PCR-RFLP 實驗提供可行的引子對。我們首先定義 T_D 為包含核苷酸 ‘A’、‘T’、‘C’ 和 ‘G’ 的模板 DNA 序列並且在此模板 DNA 序列中只有一個目標 SNP。目標 SNP 可以是 dNTP 格式，即 [dNTP1 / dNTP2]，或複合 IUPAC 碼(‘M’、‘R’、‘W’、‘S’、‘Y’、‘K’、‘V’、‘H’、‘D’、‘B’ 或 ‘N’)。在本研究中，我們不考慮缺失/插入的多型性，即 DIP 或 Indel，和多核苷酸多型性，即 MNP。 T_D 表示如下式：

$$T_D = \{B_i | i \text{ is the index of DNA sequence,} \\ \exists ! B_i \in \text{target SNP}\} \quad (1)$$

其中 B 表示鹼基， i 為 DNA 序列的位置。

在本研究中，我們也定義了 P_f 、 P_r 和 F_m ，其分別是表示前置引子、反置引子和誘發突變點位置。對於 P_f ，它必須包含目標 SNP 與在 F_m 位置上的誘發突變鹼基。此外，可用限制酶的識別序列也必須覆蓋目標 SNP 和誘發突變鹼基以辨認目標 SNP。 P_r 是從 T_D 擷取的互補且反向子序列。 P_f 和 P_r 被稱為引子對，其是 PCR-RFLP 實驗所需設計的目標。為了應用學習的概念[19-21]進行誘發突變引子設計，我們編碼學習者向量 $L_v = (F_m, F_l, P_l, R_l, \Sigma)$ 以表示這個問題的解。在 L_v 中， F_m 是如上所述的誘發突變位置，其範圍被限制在 F_{m_min} 和 F_{m_max} 之間。 F_l 是前置引子長度； P_l 是 PCR 產物長度； R_l 是反置引子長度； Σ 具有整數 0, 1, 2 或 3 的值，以分別表示 ‘A’、‘T’、‘C’ 或 ‘G’ 的誘發突變核苷酸。

2.2 誘發突變矩陣

限制酶的搜尋工作是非常乏味和耗時的，特別是使用於迭代計算方法中。為了增加限制酶搜尋效率和降低所提出方法的計算複雜性，我們在本研究中採用誘發突變矩陣(圖 1)。誘發突變矩陣可對所有假設的誘發突變核苷酸進行預先計算，以提供用於識別目標 SNP 的可用資訊。在此誘發突變矩陣中，共有 $[(F_{m_max} - F_{m_min} + 1) \times 4]$ 個元素。在 $(\text{SNP} - F_{m_max} + 1)$ 和 $(\text{SNP} - F_{m_min} + 1)$ 之間的誘發突變位置呈現在不同的列欄位中；而 ‘A’、‘T’、‘C’ 和 ‘G’ 所突變的核苷酸存在於不同的行欄位中。在此矩陣中，使用元素值 0, 1, 2 和 3 分別表示不同條件下的限制酶可用性。值 0 表示序列中的正向股和反向股都具有可用的限制酶。值 1 表示序列中的正向股或反向股僅其中一股具有可用的限制酶。值 2 表示序列中的正向股和反向股都不具有可用的限制酶。值 3 表示沒有設計誘發突變鹼基，亦沒有任何可用的限制酶。

誘發突變 矩陣	行			
	A	T	C	G
F_{m_max}	$\in N$	$\in N$	$\in N$	$\in N$
$F_{m_max}-1$	$\in N$	$\in N$	$\in N$	$\in N$
$F_{m_max}-2$	$\in N$	$\in N$	$\in N$	$\in N$
$\vdots$	$\in N$	$\in N$	$\in N$	$\in N$
$F_{m_max}-n$	$\in N$	$\in N$	$\in N$	$\in N$
F_{m_min}	$\in N$	$\in N$	$\in N$	$\in N$

$$n = F_{m_max} - F_{m_min} - 1; N = \{0, 1, 2, 3\}$$

圖 1. 誘發突變矩陣

為了更清楚的說明誘發突變矩陣，在此，我們以 $F_{m_min} = 3$ 和 $F_{m_max} = 6$ 的誘發突變位置限制範圍提出一個例子，如圖 2 所示。在矩陣中具有 0 與 1 元素的位置，表示經由誘發突變後，可以搜尋到可用的限制酶，例如當誘發突變位置為 6 時，原核苷酸 'C' 誘發突變為核苷酸 'A' 或 'T' 時，序列中的正向股和反向股皆具有可用的限制酶；而原核苷酸 'C' 誘發突變為核苷酸 'G' 時，僅序列中的正向股或反向股具有可用的限制酶。而在矩陣中具有 2 與 3 元素的位置，表示根本無可用的限制酶。

誘發突變 矩陣	行			
	A	T	C	G
6	0	0	2	1
5	3	2	3	0
4	0	2	3	3
3	1	0	1	2

圖 2. 誘發突變矩陣的例子

2.3 基於 REBASE 限制酶搜尋方法

在本研究中，一種具效率且基於 REBASE 之限制酶搜尋的新方法被使用。此方法我們提供於下列網址：

<https://sites.google.com/site/yhcheng1981/rehunt>
REBASE 中所有限制酶的識別序列皆針對目標 SNP 進行一一比對。所比對到可用的限制酶會被儲存在誘發突變矩陣中。

2.4 引子限制條件

PCR 擴增的成功率大部分由 PCR 引子限制條件所決定。引子限制條件涉及到引子長度、引子長度差、引子解鏈溫度、引子解鏈溫度差、GC 比例、GC 鉗夾、交叉二聚體、自身二聚體、髮夾結構、特異性和 PCR 產物大小 [12, 24-26]。在本研究中，我們定義引子長度介於 16nt 和 28nt 之間；引子長度差不超過 3nt；引子解鏈溫度在 45°C 和 62°C 之間；引子解鏈溫度差不超過 5°C；GC 比例介於 40% 和 60% 之間。此外，我們還評估 GC 鉗夾，並在引子中也避免了二聚體和髮夾結構，以確保引子與樣板序列的結合。最後，設計具有特異性的引子，並且讓 PCR 產物大小至少大於 100nt。

2.5 熱力學解鏈溫度計算方法

熱力學解鏈溫度計算方法用於引子設計被許多研究人員認為是較準確的。在本研究中，我們應用 SantaLucia 公式(如下公式 2)於誘發突變引子設計上。

$$Tm_{SAN}(P) = \frac{\Delta H^\circ(\text{predicted}) \times 1000}{(\Delta S^\circ(\text{salt_correction}) + R \times \ln(C_T / 4))} - 273.15 \quad (2)$$

其中 $\Delta H^\circ(\text{predicted})$ 是焓； $\Delta S^\circ(\text{salt_correction})$ 是熵校正； R 是氣體常數 (1.987 cal / Kmol)， C_T 是 DNA 濃度。

2.6 所提出的計算智慧方法

在本研究中，所提出的方法使用五個步驟用於誘發突變引子設計，流程圖如圖 3 所示。這些步驟分別是 1) 限制酶搜尋；2) 初始化族群；3) 教學過程；4) 學習過程，和 5) 終止情況判斷。底下清楚地描述這些步驟。

1) 限制酶搜尋

上述的限制酶搜尋方法和誘發突變矩陣首先用於評估可用的限制酶。我們必須知道誘發突變引子是否可用於目標 SNP。誘發突變矩陣中元素值 0 或 1 為可用，誘發突變矩陣中元素值 2 或 3 為不可用。如果誘發突變矩陣沒有任何可用的限制酶，進行誘發突變引子設計將是無意義的。

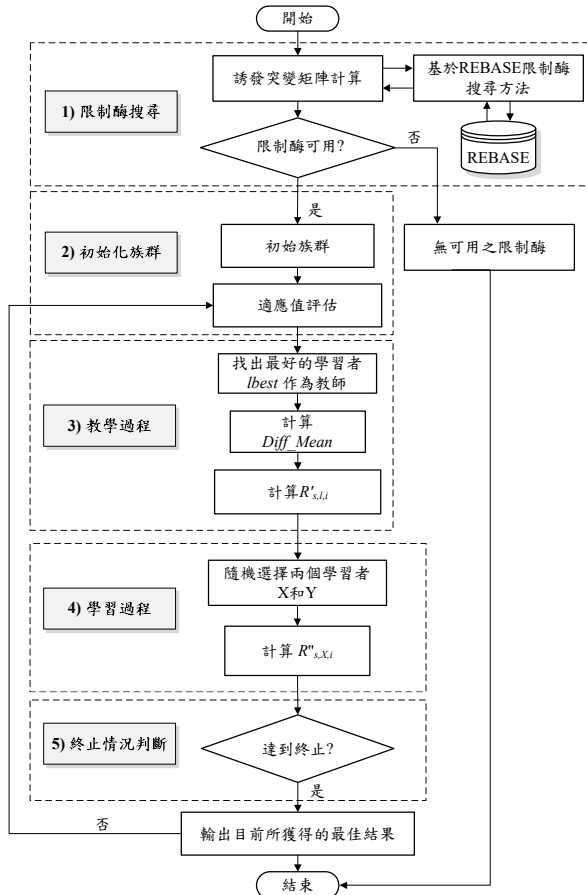


圖 3. 所提出的計算智慧方法用於誘發突變引子設計流程圖

2) 初始化族群

在所提出之以學習概念的計算智慧方法，我們以隨機產生幾十或幾百個學習者向量 $L_v = (F_m, F_l, P_l, R_l, \Sigma)$ 以提供初始解，我們稱其為初始族群。初始族群將被迭代更新，利用經驗評估函數(如下公式 3)並透過對每個學習者向量的評估來評估初始族群的適應度。在本研究方法中，我們以評估的適應值越小越好。

$$\begin{aligned} \text{Evaluation}(L_v) = & 3 \times (\text{Len}_{diff}(L_v) + \text{GC}_{proportion}(L_v) + \text{GC}_{clamp}(L_v)) \\ & + 10 \times (\text{Tm}(L_v) + \text{Tm}_{diff}(L_v) + \text{dimer}(L_v) + \\ & \text{hairpin}(L_v)) + 50 \times \text{specificity}(L_v) + \\ & 60 \times \text{product}(L_v) + 100 \times \text{RFLP}(L_v) \end{aligned} \quad (3)$$

其中 $\text{Len}_{diff}(L_v)$ 檢查前置引子和反置引子之間的長度差； $\text{GC}_{proportion}(L_v)$ 檢查前置引子和反置引子中核苷酸‘C’和‘G’的比例； $\text{GC}_{clamp}(L_v)$ 檢查出現於引子 3’末端的核苷酸是否為‘G’或‘C’； $\text{Tm}(L_v)$ 檢查前置引子和反置引子的解鏈溫度； $\text{Tm}_{diff}(L_v)$ 檢查前置引子和反置引子之間的解鏈溫度差； $\text{dimer}(L_v)$ 檢查引子的交叉二聚體和自身二聚體； $\text{hairpin}(L_v)$ 檢查引子的髮夾結構； $\text{specificity}(L_v)$ 檢查引子的特異性； $\text{product}(L_v)$

檢查 PCR 產物大小； $\text{RFLP}(L_v)$ 檢查限制酶的可用性。

3) 教學過程

在教學過程中，首先，最好的學習者被找出作為教師。考慮 s 是特定科目，並且存在 m 個科目(即， $s = 1, 2, \dots, m$)； l 是其中一位學習者，並且存在 n 個學習者(即， $l = 1, 2, \dots, n$)。在迭代 i ， $M_{s,i}$ 被表示為這些學習者在科目 s 的平均學習成果； $R_{s,l,i}$ 表示為學習者 l 在科目 s 中的學習成果； $R_{s,lbest,i}$ 表示為最好的學習者 $lbest$ 在科目 s 的學習成果； $R_{total-l,i}$ 表示為學習者 l 在所有科目的總學習成果； $R_{total-lbest,i}$ 表示為最好的學習者 $lbest$ 在所有科目的總學習成果。 $\text{Diff_Mean}_{s,l,i}$ 表示在每個科目現有平均學習結果與每個科目所相應最好的學習者學習結果之間的差異，如下式：

$$\text{Diff_Mean}_{s,l,i} = r_i \times (R_{s,lbest,i} - T_F M_{s,i}) \quad (4)$$

其中 T_F 是教學因子，其決定平均學習成果的改變， r_i 是範圍[0,1]中的隨機數。

教學因子 T_F 可由下式計算：

$$T_F = \text{round}[1 + \text{rand}(0,1)\{2-1\}] \quad (5)$$

在整個教學過程中， $R_{s,l,i}$ 由式(6)更新，更新後的值為 $R'_{s,l,i}$ 。

$$R'_{s,l,i} = R_{s,l,i} + \text{Diff_Mean}_{s,l,i} \quad (6)$$

4) 學習過程

從族群中隨機選擇兩個學習者 X 和 Y 。所選擇的學習者 X 和 Y 具有其學習結果 $R'_{total-X,i}$ 和 $R'_{total-Y,i}$ ，它們分別是 $R_{total-X,i}$ 和 $R_{total-Y,i}$ 的更新值。 $R''_{s,X,i}$ 是相互作用後的學習結果，可以通過式(7)獲得。

$$\begin{aligned} R''_{s,X,i} &= R'_{s,X,i} + r_i \times k, \\ k &= \begin{cases} R'_{s,X,i} - R'_{s,Y,i}, & \text{if } R'_{total-X,i} < R'_{total-Y,i} \\ R'_{s,Y,i} - R'_{s,X,i}, & \text{if } R'_{total-X,i} > R'_{total-Y,i} \end{cases} \end{aligned} \quad (7)$$

5) 終止情況判斷

在本研究中，當所提出的方法達到我們設定的最大迭代次數，則停止計算，並輸出目前所獲得的最佳結果。

2.7 參數設置和執行環境

在此方法中設置了兩個主要參數。一個是迭代次數，另一個是族群大小。迭代次數設置

為 1000，族群大小設置 50。執行的環境為 Intel Xeon CPU E3-1200 v2 @ 3.1 GHz x 2 和 4GB RAM 下的低階伺服器，使用 Microsoft Windows 7 專業 SP1 32 位元的作業系統。

2.8 測試資料集

從自閉症譜系障礙 (autism spectrum disorders, ASD)[27]、精神病 (psychosis)[28] 和雙極性精神失調疾病 (bipolar disorder)[29] 相關的 SLC6A4 基因篩選的 SNPs 作為測試資料集。發現共有 25 個 SNPs 具有錯配 PCR-RFLP。

3. 結果

3.1 所提出方法的執行結果

透過所提出的方法對 25 個具有錯配 PCR-RFLP 的 SNPs 進行了測試。測試結果共產生 25 對引子，總數為 50 條引子。有 18 條引子符合引子長度差的限制情況；有 31 條引子符合 GC% 的限制情況；有 19 條引子符合 GC 鉗夾；有 40 條引子符合引子解鏈溫度的限制情況；有 12 條引子符合引子解鏈溫度差的限制情況；有 50 條引子符合 PCR 產物大小的限制情況；有 25 條引子避免了交叉二聚體；有 50 條引子避免了自身二聚體；有 45 條引子避免了髮夾結構。最後，有 48 條引子符合引子特異性。結果如圖 4 所示。

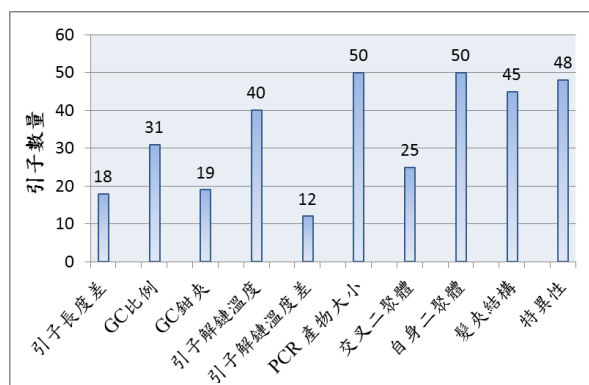


圖 4. 誘發突變引子設計結果於 25 個具有錯配 PCR-RFLP 的 SNPs

在誘發突變引子設計中，具有錯配 PCR-RFLP 的 25 個 SNPs 的引子限制的分佈顯示在圖 5~圖 10 中。設計於 17nt、20nt 和 24nt 的引子長度大於 5 條引子。所有引子長度皆設計於 16nt 至 28nt。引子長度差設計於 0 nt、2 nt

和 3nt 有 5 對引子，在 1nt 中具有 3 對引子，而其餘 9 對引子超出引子長度差設計範圍。引子解鏈溫度設計小於 45°C 有 2 條引子，大於 70°C 具有 4 條引子，其他設計皆落入設定的引子解鏈溫度範圍內。引子解鏈溫度差設計超過 5°C 具有 13 條引子，其他設計皆落入設定的引子解鏈溫度差範圍內。GC 比例設計小於 40% 具有 14 條引子，60% 至 70% 具有 4 條引子，大於 70% 具有 2 條引子，其他設計皆落入設定的 GC 比例範圍內。PCR 產物長度設計小於 100bp 具有 25 條引子對，其他設計皆落入設定 PCR 產物大小範圍內。所提出的方法平均適應值為 112.04，平均計算時間為 309.59 ms。

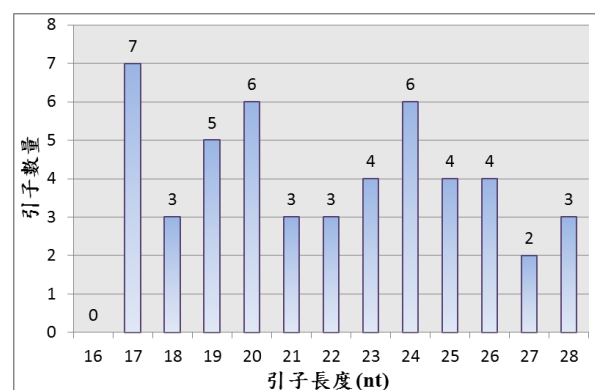


圖 5. 引子長度分佈

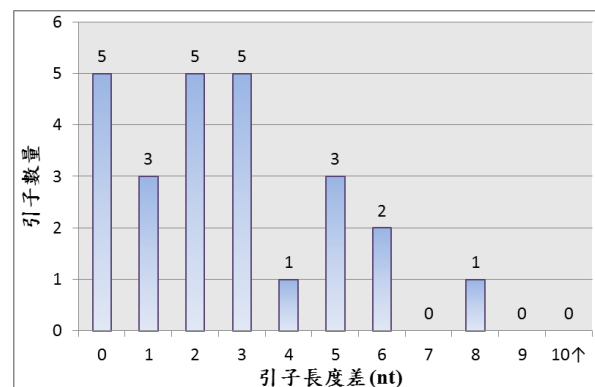


圖 6. 引子長度差分佈

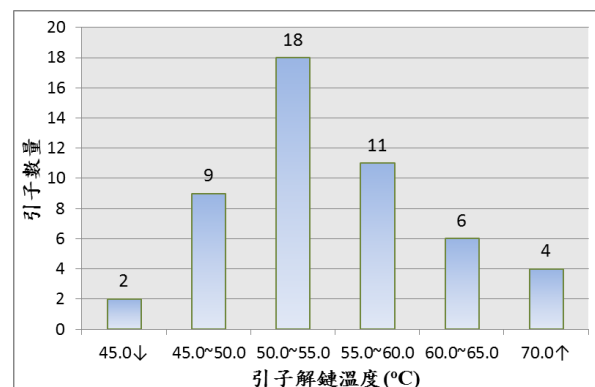


圖 7. 引子解鏈溫度分佈

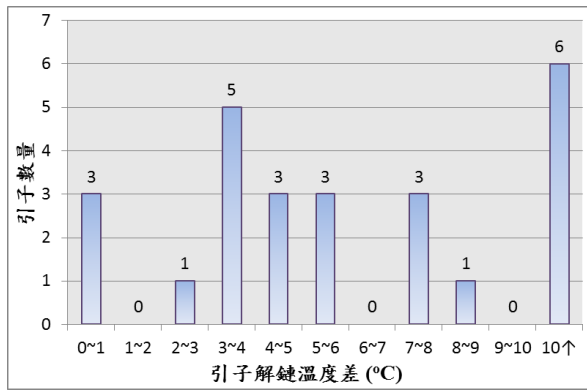


圖 8. 引子解鏈溫度差分佈

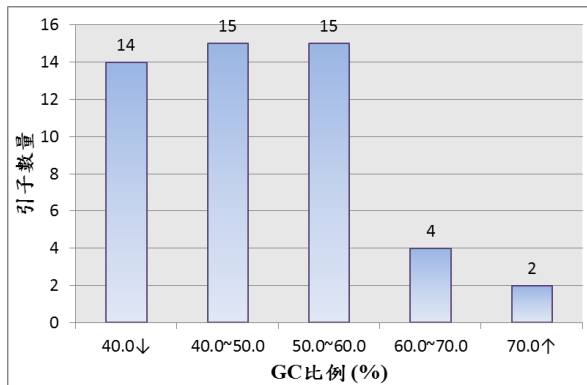


圖 9. GC 比例分佈

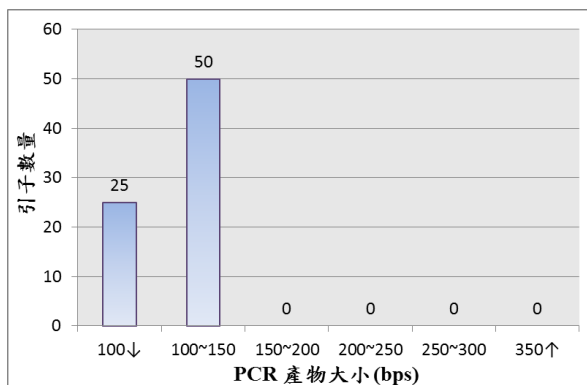


圖 10. PCR 產物大小分佈

3.2 所提出方法的優點

在 PCR-RFLP 實驗中，可用限制酶是進行成功 PCR-RFLP 的關鍵因素。然而，目標 SNP 的可用限制酶難以利用人工搜尋方式獲得。所提出的方法基於 REBASE 限制酶搜尋方法來找出可用的限制酶用於目標 SNP。因此，研究人員可以有效地獲得可用之限制酶。此外，誘發突變矩陣被引入到所提出的方法中，以在所有 REBASE 和誘發突變點中提升可用限制性酶搜尋的效率。這使得所提出的方法迅速和強健地獲得基於誘發突變點的可用限制酶。在利用計算智慧的方法中，由於初始解透過引子限

制的迭代過程被逐漸收斂到一個更好的解，因此，它可以得到更適當的誘發突變引子。較準確的熱力學解鏈溫度應用於所提出的方法也有助於 PCR-RFLP 實驗的實行。

3.3 較少的演算法參數設置

演算法特定的參數一直是使用者困擾用於解決問題的議題之一。本研究所提出的基於計算智慧的方法僅使用兩個一般參數，即迭代次數和族群大小。因此更適合於誘發突變引子設計的初學者使用。

4. 結論

在本研究中，基於計算智慧的方法結合具效率的 REBASE 之限制酶搜尋方法、誘發突變矩陣和較精確的熱力學解鏈溫度計算，以提供可用的限制酶和誘發突變引子設計。這是一種可靠和快速的方法，為目標 SNP 提供可用的限制酶，並為錯配 PCR-RFLP 實驗設計可行的誘發突變引子。本研究已經測試了使用具有錯配 PCR-RFLP 的 25 個 SNPs。結果顯示所提出的方法確實有助於設計符合條件的誘發突變引子和獲得可用的限制性酶以利目標 SNP 的 PCR-RFLP 實驗。

致謝

感謝中華民國科技部(Ministry of Science and Technology, R.O.C.)對本研究經費的補助，補助計畫編號為：MOST102-2221-E-464-004, MOST104-2221-E-464-003 和 MOST105-2221-E-324-026。

參考文獻

- [1] M. Ota, H. Fukushima, J. K. Kulski, and H. Inoko, "Single nucleotide polymorphism detection by polymerase chain reaction-restriction fragment length polymorphism," *Nat Protoc*, vol. 2, pp. 2857-64, 2007.
- [2] C. H. Yang, Y. H. Cheng, and L. Y. Chuang, "Mutagenic primer design for mismatch PCR-RFLP SNP genotyping using a genetic algorithm," *IEEE/ACM Trans Comput Biol Bioinform*, vol. 9, pp. 837-45, May-Jun 2012.
- [3] Y.-H. Cheng, L.-Y. Chuang, and C.-H.

- Yang, "Single Nucleotide Polymorphisms Polymerase Chain Reaction-Restriction Fragment Length Polymorphism Primer Design Using Genetic Algorithm with Fixed and Adaptive Stopping Criteria," *Journal of Advanced Mathematics and Applications*, vol. 2, pp. 172-181, 2013.
- [4] L.-Y. Chuang, Y.-H. Cheng, C.-H. Yang, and C.-H. Yang, "Associate PCR-RFLP assay design with SNPs based on genetic algorithm in appropriate parameters estimation," *IEEE Transactions on NanoBioscience*, vol. 12, pp. 119-127, 2013.
- [5] Y.-H. Cheng, C.-N. Kuo, and C.-M. Lai, "Effective Natural PCR-RFLP Primer Design for SNP Genotyping Using Teaching-Learning-Based Optimization with Elite Strategy," *IEEE Transactions on NanoBioscience*, 2016.
- [6] R. Zhang, Z. Zhu, H. Zhu, T. Nguyen, F. Yao, K. Xia, D. Liang, and C. Liu, "SNP Cutter: a comprehensive tool for SNP PCR-RFLP assay design," *Nucleic Acids Res*, vol. 33, pp. W489-92, Jul 1 2005.
- [7] L. Y. Chuang, C. H. Yang, Y. H. Cheng, D. L. Gu, P. L. Chang, K. H. Tsui, and H. W. Chang, "V-MitoSNP: visualization of human mitochondrial SNPs," *BMC Bioinformatics*, vol. 7, p. 379, 2006.
- [8] H. W. Chang, L. Y. Chuang, Y. H. Cheng, Y. C. Hung, C. H. Wen, D. L. Gu, and C. H. Yang, "Prim-SNPing: a primer designer for cost-effective SNP genotyping," *Biotechniques*, vol. 46, pp. 421-31, May 2009.
- [9] R. B. Wallace, J. Shaffer, R. F. Murphy, J. Bonner, T. Hirose, and K. Itakura, "Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: the effect of single base pair mismatch," *Nucleic Acids Res*, vol. 6, pp. 3543-57, Aug 10 1979.
- [10] C. W. Dieffenbach, T. M. Lowe, and G. S. Dveksler, "General concepts for PCR primer design," *PCR Methods Appl*, vol. 3, pp. S30-7, Dec 1993.
- [11] K. A. Abd-Elsalam, "Bioinformatic tools and guideline for PCR primer design," *african Journal of biotechnology*, vol. 2, pp. 91-95, 2003.
- [12] L. Y. Chuang, Y. H. Cheng, and C. H. Yang, "Specific primer design for the polymerase chain reaction," *Biotechnol Lett*, vol. 35, pp. 1541-9, Oct 2013.
- [13] J. SantaLucia, Jr., "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics," *Proc Natl Acad Sci U S A*, vol. 95, pp. 1460-5, Feb 17 1998.
- [14] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen, "Primer3--new capabilities and interfaces," *Nucleic Acids Res*, vol. 40, p. e115, Aug 2012.
- [15] L. C. Li and R. Dahiya, "MethPrimer: designing primers for methylation PCRs," *Bioinformatics*, vol. 18, pp. 1427-31, Nov 2002.
- [16] F. M. You, N. Huo, Y. Q. Gu, M. C. Luo, Y. Ma, D. Hane, G. R. Lazo, J. Dvorak, and O. D. Anderson, "BatchPrimer3: a high throughput web application for PCR and sequencing primer design," *BMC Bioinformatics*, vol. 9, p. 253, 2008.
- [17] S. Arvidsson, M. Kwasniewski, D. M. Riano-Pachon, and B. Mueller-Roeber, "QuantPrime--a flexible tool for reliable high-throughput primer design for quantitative PCR," *BMC Bioinformatics*, vol. 9, p. 465, 2008.
- [18] J. Ye, G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden, "Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction," *BMC Bioinformatics*, vol. 13, p. 134, 2012.
- [19] Y.-H. Cheng, "Computational intelligence-based polymerase chain reaction primer selection based on a novel teaching-learning-based optimisation," *IET Nanobiotechnology*, vol. 8, pp. 238-246, 2014.
- [20] Y.-H. Cheng, "Estimation of Teaching-Learning-Based Optimization Primer Design Using Regression Analysis for Different Melting Temperature Calculations," *IEEE Transactions on NanoBioscience*, vol. 14, pp. 3-12, 2015.
- [21] Y.-H. Cheng, "A novel teaching-learning-based optimization for improved mutagenic primer design in mismatch PCR-RFLP SNP genotyping," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 13, pp. 86-98, 2016.
- [22] A. Haliassos, J. C. Chomel, S. Grandjouan, J. Kruh, J. C. Kaplan, and A.

- Kitzis, "Detection of minority point mutations by modified PCR technique: a new approach for a sensitive diagnosis of tumor-progression markers," *Nucleic Acids Res*, vol. 17, pp. 8093-9, Oct 25 1989.
- [23] A. Haliassos, J. C. Chomel, L. Tesson, M. Baudis, J. Kruh, J. C. Kaplan, and A. Kitzis, "Modification of enzymatically amplified DNA for the detection of point mutations," *Nucleic Acids Res*, vol. 17, p. 3606, May 11 1989.
- [24] M. J. McPherson, G. R. Taylor, and P. Quirke, *PCR, a practical approach*: Oxford University Press, USA, 1991.
- [25] S. Rozen and H. Skaletsky, "Primer3 on the WWW for general users and for biologist programmers," *Methods in Molecular Biology*, vol. 132, pp. 365-386, 2000.
- [26] J. Sambrook and D. W. Russell, "Molecular cloning: a laboratory manual," *Cold Spring Harbor Laboratory Press*, vol. 2, pp. 8.1-8.126, 2001.
- [27] T. Sakurai, J. Reichert, E. J. Hoffman, G. Cai, H. B. Jones, M. Faham, and J. D. Buxbaum, "A large-scale screen for coding variants in SERT/SLC6A4 in autism spectrum disorders," *Autism Res*, vol. 1, pp. 251-7, Aug 2008.
- [28] T. E. Goldberg, R. Kotov, A. T. Lee, P. K. Gregersen, T. Lencz, E. Bromet, and A. K. Malhotra, "The serotonin transporter gene and disease modification in psychosis: Evidence for systematic differences in allelic directionality at the 5-HTTLPR locus," *Schizophr Res*, vol. 111, pp. 103-8, Jun 2009.
- [29] L. Mandelli, M. Mazza, G. Martinotti, M. Di Nicola, D. Tavian, E. Colombo, S. Missaglia, D. De Ronchi, R. Colombo, and L. Janiri, "Harm avoidance moderates the influence of serotonin transporter gene variants on treatment outcome in bipolar patients," *J Affect Disord*, vol. 119, pp. 205-209, 2009.