

生物資訊學在植物病害診斷實務上之應用

胡仲祺 助理教授¹ 張世忠 技正²

1 國立中興大學生物科技學研究所

電子郵件：cchu@dragon.nchu.edu.tw; 傳真：04-22855720

2 行政院農業委員會動植物防疫檢疫局

摘 要

生物資訊學技術為植物病理學基礎與應用研究上所不可或缺的工具。病原之不同株系在被引入新地區並入侵新寄主後,可經由基因體之突變(mutation)、重組(recombination)、或假重組(re-assortment)而產生新株系並造成嚴重的新流行病發生。而許多國家也以此為由,對他國之農產品設下檢疫關卡。我國防檢疫單位除需面對各國對於不同植物病原株系所設之規定外,尚需設法防堵國外病原株系之入侵,以保護本國之農產品。同時也需要具體資料以追蹤境外病毒之入侵感染途徑、傳播範圍、及其是否已與本土病毒發生基因重組。而這些工作都需要依賴生物資訊學技術之應用,協助分析比對病原基因體序列之特性,方能迅速提出具體之分子證據,追蹤病原入侵途徑,判定病原是否屬於本土已有之株系。為因應加入 WTO 後,不同地區植物病毒大量流通所帶來的衝擊,行政院農委會動植物防疫檢疫局已建置台灣地區植物病毒生物資訊資料庫與通報介面,提供中文化的操作介面與簡單常用的生物資訊學技術之簡介,期望能作為防止境外植物病原入侵、追蹤境外病原之基本工具。除可協助政府相關單位隨時掌控疫情資料外,更可作為貿易諮商談判時之具體籌碼。

關鍵詞：生物資訊學技術、資料庫、植物病毒、基因體序列分析

緒 言

電腦與資訊科技已與大眾之日常生活密不可分。而植物病理學發展至今，亦已與資訊科學密不可分。在許多植物病原之全基因體序列陸續被解出後，植物病理學家也獲得前所未有的利器，以了解植物病原何以造成植物病害之分子機制、其致病基因之表現與調控、及其與寄主植物和環境間交互作用。然而，隨著大量的核酸與蛋白質序列資訊迅速累積，人腦已經無法有效處理、分析、並擷取有用的知識，而必須仰賴先進之電腦與資訊技術輔佐。生物資訊學技術也因此而扮演著日益重要的角色。

生物資訊學廣義的定義即為使用電腦或資訊學技術從事生物科學方面的研究，包括實驗數據分析、文獻資料查詢、以至撰寫論文或使用電腦繪製圖形。因此，生物資訊學嚴格來說並不能算是一門新的學門。而生物資訊學狹義的定義，則依據舊金山州立大學Dr. R. L. Bernstein認為是使用電腦工具去儲存、應用、分析核酸與胺基酸序列及蛋白質結構資料。目前的生物資訊學則已經可區分為三個不同的研究方向(Baxevanis and Ouellette, 1998; Brenner et al., 1998)：

1. 生物統計與演算法之研發：此部份由生物統計學家與數學家等之研究為主，藉由目前已知之生物巨分子序列資料及其已確定知生物功能，以各種生物統計法分析歸納出序列與功能之間的關係，或發展出新的統計方法或演算式(algorithm)以探討目前方法所無法分析之關聯性；例如在資料庫的搜尋與基因序列比對的研究上，由最早的全域比對演算法(global alignment algorithm)逐漸演變出區域比對演算法(local alignment algorithm)，再演變為加入統計分析之期望值門檻(E-value cut-off)與極端值分布統計(extreme value distribution)，而成為目前最快速且流行的BLAST工具。

2. 電腦軟體工具之研發：此部份研究以資訊學家與生物學家之共同研究為主，主要為利用上述第一項領域所研發之統計結果與新演算

式作基礎，開發出適合某種特定電腦平台(或是合所有平台)之軟體工具，特別著重於使用者介面之設計，以符合”友善”的要求。例如目前在學術界廣泛使用的GCG package與Vector NTI等軟體，即為整合了許多不同研究學者所撰寫的單一目的程式，賦予相同的指令組與使用者介面(分別針對UNIX與Windows操作系統)，而成為一般使用者皆能輕易上手的套裝軟體組。

3. 生物巨分子資料之分析與解譯：此部份研究以生物學家為主，包括應用上述二項領域所產生之工具以分析並解譯經由實際進行的生物實驗所產生的結果、以及將所分析解譯後所歸納的知識再交由上述二項領域的專家研究以產生更新更好的演算法與軟體工具。例如由人類全部基因體的解序完成，我們可以利用各種電腦分析方法與軟體工具歸納出某種基因與某種表現性狀之間的關聯性。而這種關聯性的假說，又可以提供給生物學家設計一連串的具體實驗工作加以驗證。

上述三個次領域的研究環環相扣，相輔相成，不斷的將生物資訊學的基礎研究與應用層面帶向更高的境界。而隨著電腦、資訊、與分子生物科技的不斷進步，生物資訊學工具也將在防疫檢疫的工作方面扮演關鍵性的角色。本文將針對生物資訊學技術在植物病理學方面的應用進行簡介並以行政院農業委員會動植物防疫檢疫局所建置之植物病毒生物資訊資料庫為例說明從事防檢疫工作相關人員如何應用此類資料庫輔助本身工作。期能妥善的運用各種網路資源與基礎生物資訊工具，以利防疫檢疫工作之推動。

生物資訊學在植物病理學研究之應用實例

1. 生物資訊學技術在病原真菌與線蟲方面的應用：

目前在GenBank中已有許多植物病原真菌與線蟲的部分基因體序列(Benson et al., 2004)。而維吉尼亞生物資訊研究所的Dr. Tyler即利用

序列分析比對等生物資訊學分析方法研究*Phytophthora*屬真菌與其寄主辨識的分子機制，並已成功找出至少一個由*Phytophthora* 所編碼的Avr基因可在帶有Rps1b基因的大豆寄主上引起系統性防禦反應(Tyler, 2002)。在線蟲方面，許多研究人員利用18S與28S ribosome 基因、粒線體基因、或ITS序列輔助進行線蟲的分類，預期在不久的將來可將線蟲的基因體標誌儲存於資料庫中，成為每一種線蟲特殊的”條碼”，以粒線蟲鑑定工作的進行(Powers, 2004)。

2. 生物資訊學技術在病原細菌方面的應用：

在病原細菌方面的應用，則由於數個植物病原細菌已經完全解序，研究人員得以對於病原細菌的基因體密碼做分子層次的了解。例如 Genin 與 Boucher(2004) 利用已經全部解序完畢之*Ralstonia solanacearum* GMI1000基因體作為材料，研究此細菌如何演化以適應不同寄主植物，以及其必要的致病基因(如第三型分泌系統基因群)如何調控等分子機制。未來對於病原細菌的分類鑑定、抗病育種、病害管理等工作將具有極大的貢獻。

3. 生物資訊學技術在植物病毒方面的應用：

由於病毒的基因體極小，較容易進行全基因體解序工作，目前的基因庫中所包含的病毒全基因體資料為所有生物中最多者。因此，生物資訊學技術在病毒之應用也最為普遍。已有許多研究利用全基因體序列資料的比對，指出來自不同地理區域之不同病毒株系(strain)，可經由其基因體之重組而產生新株系，並可利用基因體標記追蹤病毒的入侵可能途徑。而新病毒的產生通常可造成新流行病之發生。例如在雙生病毒中，已有烏干達地區發現樹薯嵌紋病毒的重組株，造成嚴重損失，在印度也發現單基因體之樹薯嵌紋病毒經由假重組而成為雙基因體；另有番茄捲葉病毒在西班牙地區發生基因重組而造成寄主範圍增大與粉蝨傳播能力增加的狀況(Harrison and Robinson, 1999; Fondong et al., 2000; Monci et al., 2002; Saunders et al., 2001; 2002等)。這些研究

皆利用本土建構之資料庫，以生物資訊學技術分析比對不同來源的病毒基因體，而推斷出新病毒的起源與可能的入侵途徑。

台灣地區植物病毒生物資訊資料庫之建立與應用

台灣地處亞熱帶及熱帶氣候區，天然環境適合栽培種類繁多之農作物，加上農民先進之栽培技術，使得農產品具有優良的品質，極具有國際競爭力。我國在加入世界貿易組織（World Trade Organization, WTO）後，農產品進出口量逐年增加，致使台灣成為外來危險性有害生物入侵之高危險地區，再加上一年四季均可栽植作物，媒介昆蟲存活容易，也極為適合植物病毒病害發生與傳播。由於目前並無有效之化學藥劑可治療植物病毒病害，因此對於植物病毒之防治工作，首重預防，特別是防止境外感染原之入侵。我國防檢疫單位除需面對各國對於不同植物病毒所設之檢疫關卡外，尚需設法防堵國外病毒藉由進口農產品潛伏感染而入侵，以保護本國之農作物，同時也需要具體資料以追蹤境外病毒之入侵感染途徑、傳播範圍、及其是否已與本土病毒發生基因重組。而這些工作都需要依賴本土植物病毒基因體資料庫與通報介面之建立，方能迅速提出具體證據，判定病毒是否屬於本土已有之株系，以作為檢疫依據及貿易諮商談判之籌碼。

因此，由行政院農業委員會動植物防疫檢疫局企劃組主導構建本資料庫。主要目的為提供各階層的決策者與工作人員必要的相關資訊，以因應國際貿易大幅增加後，世界各地區不同病毒隨農產品或其他商品互相傳播或進行基因重組所衍生的各種植物防疫檢疫上的問題，追蹤病毒入侵途徑，掌控疫情資料。

同時，亦提供具有中文介面的生物資訊學研究工具與淺顯易懂的生物資訊學理論介紹，可提供國內研究人員了解並進行基本的核酸或胺基酸序列搜尋比對等工作，並具有與國際接軌且國內獨立的病毒資

料庫，在國外資料庫暫停服務或網路塞車時，仍能快速進行各種研究工作。

此外，為掌握防檢疫工作之時效，儘速防堵有害病毒入侵，網站內提供中文化之新病毒通報系統，供第一線工作人員與各級決策單位迅速掌握疫情，採取對策。同時並提供討論區，讓相關工作人員能自由發表意見，交換工作經驗，解決所遭遇之難題，並進行腦力激盪而開發創新。

在防檢局企劃組及各位先進之支援下，本資料庫希望能為大家提供最新的資訊與完整的服務，作為工作時最方便的基本工具。

本站的資訊系統架構如圖一，其內容與使用方法簡述於後：

1. 台灣地區植物病毒生物資訊庫網站：

為讓工作人員可輕易使用此資料庫，目前使用最方便之網路介面，利用PHP與MySQL架構後端資料庫與前端使用者介面，讓使用者僅需藉由滑鼠的點選，或填寫簡單的表單，即可獲得所欲知之資訊，如圖二所示，網站首頁即包含通往其他資源與網頁之超連結。另為顧及網站與資料庫之安全性，使用者必須先行申請帳號與密碼，才可進入資料庫。如需使用，以電子郵件註明使用目的、所屬機關單位或個人、聯絡方式等相關資料後，傳送至 cchu@dragon.nchu.edu.tw 以獲得帳號與密碼。

2. 中文化病毒核酸與蛋白質序列比對分析介面：

在使用專業資料庫時，大部分第一線工作人員所遭遇的困擾為語言問題，當在實驗室中解出所研究病毒之核酸序列後，面對NCBI或EMBL資料庫之搜尋介面，常有無從著手的迷惑。因此目前所建構之最常用基因體資料庫GenBank之中文使用者介面如圖三，可方便不熟悉英文的使用者也能輕易搜尋與比對GenBank資料庫。但因該資料庫中的內容仍為英文，故本資料庫網站亦提供中文說明，解釋各關鍵參數的涵義（如bit score, E-value）等，方便使用者解釋自己的搜尋結果。

3. 植物病毒基因體中文化資料庫：

資料庫中目前已收集台灣地區植物病毒資料170筆，並有一千筆以上之自GenBank中蒐集之世界其他地區病毒核酸序列資料。目前以MySQL資料庫軟體將上述所蒐集之資料重新整理，由目前在UNIX與LINUX操作系統中與網路介面中較為流行的PHP語法介面進行資料庫管理與維護，以建立本土化之植物病毒資料庫。同時具有中文介面以利第一線工作人員查詢比對。(如圖四)

4. 中文化通報介面：

中文化植物病毒通報介面可以供第一線工作人員於發現外來病毒株系入侵後，迅速進行通報作業，並可作為管理階層即時掌控資訊之基本工具。通報介面欄位包括：發現地點、發現時間、寄主植物、病徵、核酸序列資料、蛋白質序列資料、結構蛋白分子量資料、血清反應等。並可定期統計不同病毒之發生頻度及狀況，進行與其他地區相關病毒之比對分析(如圖五)。

5. 國外期刊中台灣地區與大陸地區植物病毒相關之文獻資料：

除了提供使用者利用核酸或胺基酸序列進行比對之外，本站也提供國外期刊台灣地區與大陸地區植物病毒文獻資料的查詢。此功能是将所有已知的植物病毒按照科別分群，並以其屬名為代表呈現。每個病毒之屬名皆提供超連結至美國National Center for Biotechnology Information所維護的PubMed醫學文獻資料庫，可進行自動化查詢。內建的查詢分別為『" virus genus" AND Taiwan[AD]』與『(" virus genus" AND China[AD]) NOT Taiwan[AD]』。使用者僅需點選該病毒屬名，即可自動獲得國外期刊中所發表有關台灣或大陸植物病毒的最新消息與文獻。隨時追蹤自己正在研究的相關病毒資訊，掌握最新動態。(圖六)

6. 其他輔助功能(討論區圖七)：

為使從事防疫檢疫相關工作的同好們能隨時交換工作經驗、傳遞訊息、與聯絡感情，本站提供二個討論區(生物資訊與植物病毒)供大家隨時發表高見。本資料庫並將常用的植物病毒研究方法(例如蔗糖密度梯度之計算，如圖七)整合成網路介面，方便使用者隨時上網查詢。

結 語

此植物病毒生物資訊庫可提供防檢疫相關工作人員方便使用之操作介面及自動統計、比對系統，可迅速供給資訊以判定病毒是否屬於台灣地區本土病毒，不受限於國際網路之流量與頻寬。預期此資料庫及通報系統將可以協助政府相關單位隨時掌控疫情資料，並作為防止境外植物病毒入侵、追蹤境外病毒、或進行貿易諮商談判之基本工具。未來將持續擴充台灣地區之植物病毒基因體資料庫，並擬針對中國大陸、東南亞地區國家、與美國等重要農產品輸出國分別建立次資料庫(Sub-database)，加速搜尋與分析速度。並針對使用者之反應改善使用者介面，同時進行資料庫系統與資料之維護作業自動化，同時將此資料庫平台推廣應用於動物病毒基因體之研究方面，以協助政府單位解決進出口貿易上防疫檢疫相關問題。

引用文獻

1. Baxevanis, A. D. and Ouellette, B. F. F. 1998. Bioinformatics: A practical guide to the analysis of genes and proteins. John Wiley and Sons, Inc. New York.
2. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. (2004). GenBank: update. *Nucleic Acids Research* 32, D23-D26.
3. Brenner, S., Lewitter, F., Patterson, M., and Handel, M. 1998. *Trends guide to bioinformatics*. Elsevier Science.
4. Fondong, V. N., Pita, J. S., Rey, M. E., de Kochko, A., Beachy, R. N. & Fauquet, C. M. (2000). Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *J. Gen. Virol.* 81, 287-97.
5. Fraile, A., Alonso-Prodos, J. L., Aranda, M. A., Bernal, J. J., Malpica, J. M. & Garcia-Arenal, F. (1997). Genetic exchange by recombination or

- reassortment is infrequent in natural population of a tripartite RNA plant virus. *J. Virol.* 71, 934- 940.
6. Genin, S. & Boucher, C. 2004. Lessons learned from the genome analysis of *Ralstonia solanacearum*. *Annu. Rev. Phytopathol.* 42:107–34.
 7. Harrison, B. & Robinson, D. (1999). Natural Genomic and Antigenic Variation In Whitefly-Transmitted Geminiviruses (Begomoviruses). *Annu. Rev. Phytopathol.* 37, 369-398.
 8. Hellwald, K. H. & Palukaitis, P. (1995). Viral RNA as a potential target for two independent mechanisms of replicase- mediated resistance against cucumber mosaic virus. *Cell* 83, 937-46.
 9. Monci, F., Sanchez-Campos, S., Navas-Castillo, J. & Moriones, E. (2002). A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. *Virology* 303, 317-26.
 10. Powers, T. 2004. Nematode molecular diagnostics: from bands to barcodes. *Ann. Rev. Phytopath.* 42: 367-383.
 11. Saunders, K., Bedford, I. D. & Stanley, J. (2001). Pathogenicity of a natural recombinant associated with ageratum yellow vein disease: implications for geminivirus evolution and disease aetiology. *Virology* 282, 38-47.
 12. Saunders, K., Salim, N., Mali, V. R., Malathi, V. G., Briddon, R., Markham, P. G. & Stanley, J. (2002). Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: evidence for acquisition of a DNA B component by a monopartite begomovirus. *Virology* 293, 63-74.
 13. Tyler, B. 2002. Molecular basis of recognition between *Phytophthora* pathogens and their hosts. *Annu. Rev. Phytopathol.* 40:137–67.

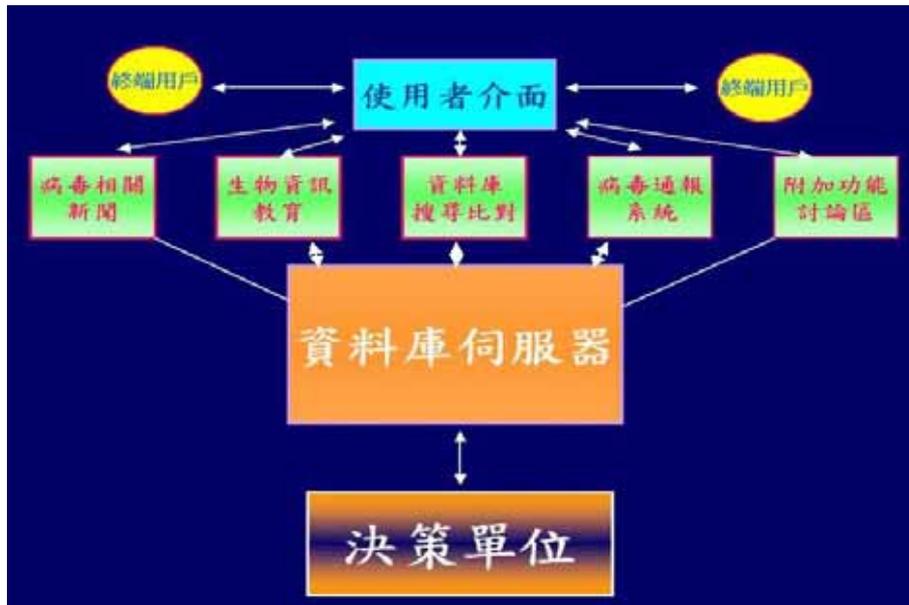
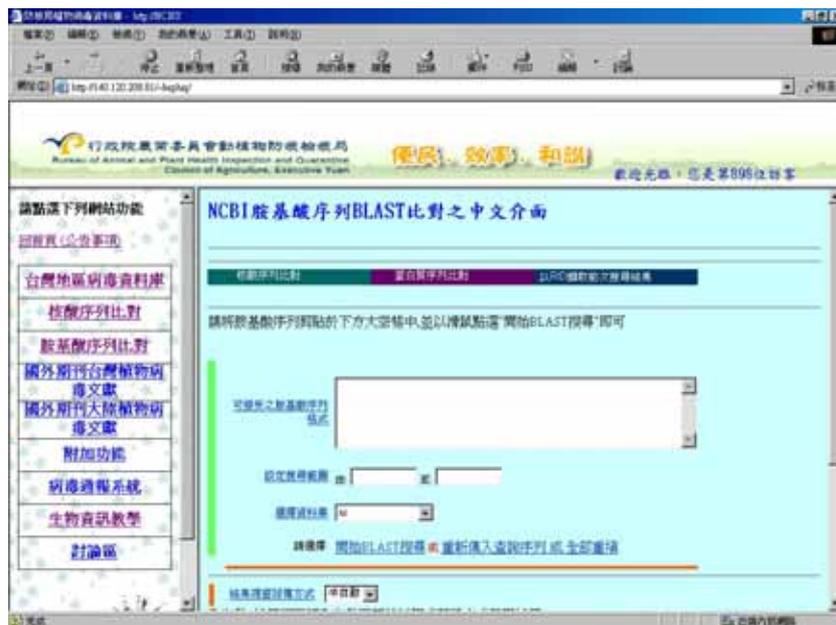
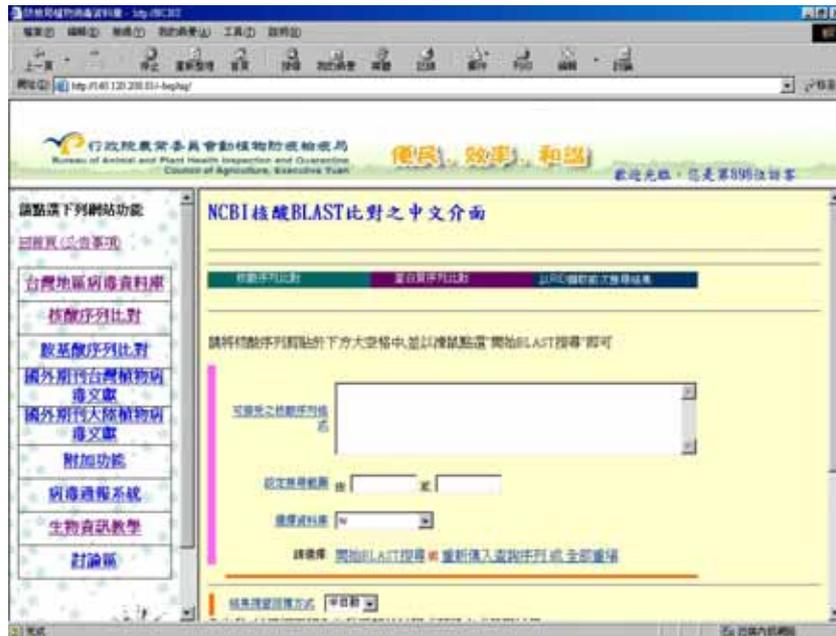


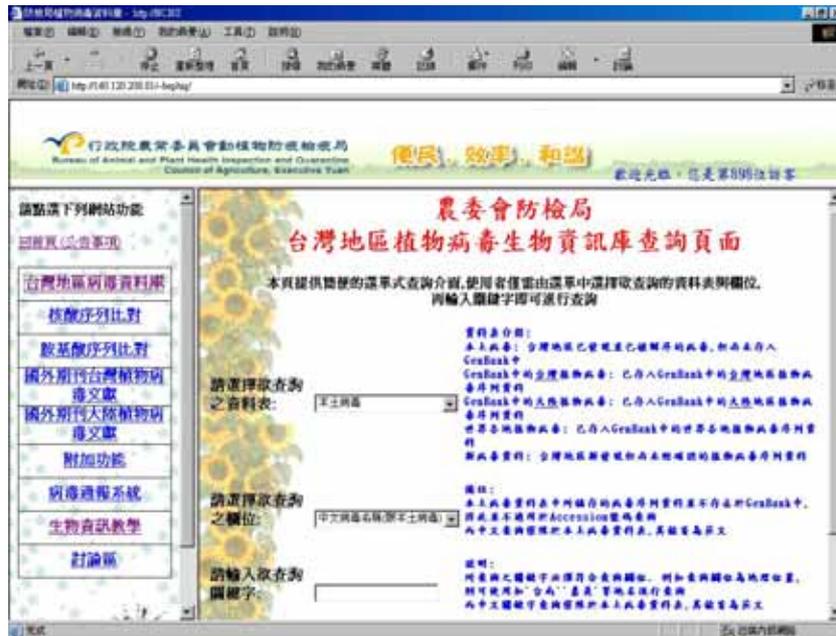
圖1、植物病毒生物資訊庫整體架構



圖2、植物病毒生物資訊庫網站首頁，包含所有訊息與資源所必須之超連結。



圖三、NCBI-BLAST核酸與胺基酸序列搜尋比對中文化使用者介面。



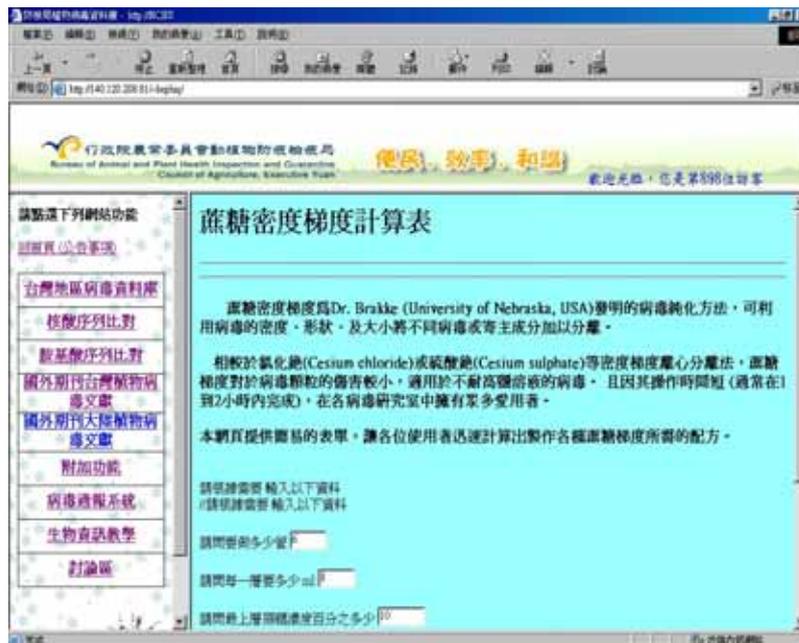
圖四、植物病毒基因體中文化資料庫查詢頁面(左)與查詢結果範例(右)。



圖五、新發現入侵病毒通報介面。



圖六、NCBI-PubMed文獻資料庫之簡易搜尋介面。



圖七、討論區(左)與蔗糖密度梯度計算表使用者介面。