

受稻熱病菌誘導之水稻表現序列標幟 (EST) 的比較分析¹

呂秀英^{2,4} 邱怡嘉³ 陳政道² 魏夢麗² 呂椿棠²

摘 要

呂秀英、邱怡嘉、陳政道、魏夢麗、呂椿棠。2007。受稻熱病菌誘導之水稻表現序列標幟 (EST) 的比較分析。台灣農業研究 56:261-280。

稻熱病由真菌 (*Magnaporthe grisea* (Hebert) Barr) 所引起，是嚴重影響水稻 (*Oryza sativa* L.) 生產的世界性重要病害。表現序列標幟 (expressed sequence tag, EST) 是 cDNA 序列片段，直接研究 EST 可獲取基因表現訊息。為探討水稻不同品種於不同植株部位、不同發育時期感染稻熱病之不同功能基因的表現差異特性，本研究擷取 NCBI 的 dbEST 資料庫中不同水稻品種、植株部位及發育時期的受稻熱病菌誘導之水稻 EST 資料，共 14 個 cDNA 基因庫，就各基因庫分別進行 EST 的功能註解、分類和表現量統計，並藉由集群分析和對應分析檢測不同基因庫間之基因表現型式的差異性及表現基因的功能類別。結果顯示，野生稻受稻熱病菌侵染後，代謝、細胞生長和形態發生、次細胞定位和傳輸促進作用等相關功能的基因表現量，高於秈稻和粳稻；水稻葉片在不同生長階段中感染稻熱病，其表現出的能量和代謝功能之基因表現量以成熟期為最高，其次為分蘖期，苗期則最低；受稻熱病菌誘導的水稻 EST 序列若混合了不同品種及不同部位組織，與使用單一品種及固定部位組織在基因表現上有明顯差異。這些發現，可提供了解水稻對稻熱病防禦機制的參考。

關鍵詞：水稻、稻熱病、表現序列標幟、功能註解及分類、基因表現量統計、集群分析、對應分析。

前 言

廣泛應用資料庫中的表現序列標幟 (expressed sequence tag, EST)，可用來探討不同組織的基因表現 (Adams *et al.* 1991; Rudd 2003; Wei *et al.* 2005a, b)。EST 是將某特定生長發育期或某特定生理狀態下之組織所表現的訊息 RNA (messenger RNA, mRNA) 經反轉錄酶 (reverse transcriptase) 產生互補 DNA (complementary DNA, cDNA) 後，自其 5' 端進行定序所得到的序列稱之，通常長度是 200-500 bp (Adams *et al.* 1991)。EST 雖然只是 cDNA 的片段序列，但具快速且低廉的優點，故成為連接結構基因組學和功能基因組學的一道橋樑，進行 EST 分析可提供基因辨識、新基因發現及基因表現差異性等研究，是最快速獲得序列資訊及提供功能性基因組研究之來源 (Adams *et al.* 1991,

1. 行政院農業委員會農業試驗所研究報告第 2301 號。接受日期：96 年 11 月 29 日。
2. 本所作物組研究員、計畫助理、助理研究員及助理研究員。台灣 台中縣 霧峰鄉。
3. 私立朝陽科技大學生物技術研究所碩士。台灣 台中縣 霧峰鄉。
4. 通訊作者，電子郵件：iying@wufeng.tari.gov.tw；傳真機：(04)23390528。

1992; Polymeropoulos 1993; Bonaldo *et al.* 1996; Rounsley *et al.* 1996; Ewing *et al.* 1999; Jongeneel 2000; Kantety *et al.* 2002; Rudd 2003; Wei *et al.* 2005a, b)。隨著「國際水稻基因組定序計畫」(IRGSP) 於 2002 年 12 月宣佈水稻 (*Oryza sativa* L.) 基因組定序草圖完成 (Feng *et al.* 2002; Sasaki *et al.* 2002; Tyagi *et al.* 2004; IRGSP 2005)，對水稻功能基因組的研究準確度將有大幅提昇 (Wei *et al.* 2005a)。針對水稻受病蟲害或逆境誘導之 EST 進行功能註解 (functional annotation)，將有助於瞭解水稻對病蟲害、逆境壓力之防禦機制的研究。

稻熱病 (古稱稻瘟)，由真菌 (*Magnaporthe grisea* (Hebert) Barr) 所引起，是影響水稻生產最嚴重的世界性真菌性病害 (Liu *et al.* 1981; Liang 2005)。根據研究報告指出，1975 至 1990 年間因稻熱病所導致的全球稻作糧食損失平均每年高達約一千萬噸 (Baker *et al.* 1997)。稻熱病自水稻秧苗期至成熟期之整個生長發育時期及稻株各個部位 (苗、葉、葉舌、節、穗) 皆會被感染發病，病徵因各部位發病的病徵稍有差異，一般病徵為發病初期先於葉面上形成褐色或暗綠色小斑點，如環境適合，擴大成紡錘形，此時病斑周圍呈黃色，中間赤褐色，內部灰白色；嚴重時葉片枯萎甚至全株枯死。稻熱病菌遺傳上的多樣性和複雜性是稻熱病廣泛流行的主要原因，目前生產上對稻熱病的防治並無特效的方法，多是依賴水稻品種自身的抗性來抵禦病害的發生，因此稻熱病抗性的遺傳和育種研究就具有十分重要的意義 (Liang 2005)。目前國內外對基因層次對稻熱病的抗病機制研究上，多集中在抗病基因的鑑定和選殖上，但對於水稻防禦稻熱病菌侵染的基因表現差異，迄今相關研究仍相當有限。例如 Rauyaree *et al.* (2001) 以稻熱病菌感染 48 小時的水稻葉片組織 EST，經過聚集 (clustering) 及接合 (assembly) 之常規化處理以減少多餘性 (redundancy) 後所產生的接合序列 (contig 或 consensus sequence)，經由資料庫比對進行功能註解及分類，結果得知以表現逆境防禦反應之基因量最多 (21%)，其他則是參與代謝、蛋白質合成、細胞訊息傳遞等相關功能的基因，但也有 19% 尚未清楚是何種功能基因。Jantasuriyarat *et al.* (2005) 利用抗病、部分抗病及感病之不同來源的水稻葉片組織 EST，檢測其接種稻熱病菌 6 及 24 小時後的基因功能表現，透過 EOG (Eukaryotic Orthologous Group) 進行功能註解及分類，發現其中只有 42% 可推測出其功能基因，而且防禦、訊息傳遞機制、細胞循環控制、細胞分裂等相關功能的基因表現量會增加，以抵抗病原菌之侵害。

植物與病原相互作用的過程中一般涉及病原與宿主間的相互識別、訊號傳導、植物抗性基因活化等一系列的生理生化反應，從不同水稻品種於不同植株部位、不同發育時期感染稻熱病之基因表現差異的角度進行研究，結合 EST 的生物資訊學分析，以獲取基因表現豐富度的變化數據，可更有助於了解水稻受稻熱病菌侵染後水稻防禦反應體系的構成狀況。至 2007 年 7 月 31 日止，登錄在 NCBI 的 dbEST 資料庫中的水稻 EST 序列數量已累積近 1,270,889 萬條，其中水稻受重要病蟲害所誘導之 EST 中以受稻熱病菌誘導之序列資料為最多，高達 84,705 條。因此，本研究從 dbEST 資料庫中擷取受稻熱病菌誘導之水稻 EST 資料，篩選出定序材料分別取自不同品種於不同植株部位、不同發育時期的水稻組織，共 14 個 cDNA 基因庫，進行基因功能的註解、分類及表現量統計，然後利用集群分析 (clustering analysis) 及對應分析 (correspondence analysis) 之多變數統計技術 (multivariate analysis)，以探討水稻不同品種於不同植株部位、不同發育時期感染稻熱病之不同功能基因的表現差異特性。

材料與方法

試驗材料

本研究擷取 NCBI 之 dbEST 資料庫中受稻熱病菌誘導所產生之 84,705 條水稻 EST 序列，依登錄時間由遠至近排序，分屬於 14 個 cDNA 基因庫，本研究各以 BF-1、BI-2、BI-1、BM-2、BM-1、AF、BQ、BU、CBir36、CBbare、BF-2、CD、CX 及 DN 之 14 個代號來表示，並下載儲存於不同的 FASTA 格式之序列檔案，其相關資訊整理如表 1。這些 EST 分別取自於不同品種(包含野生種)、不同植株部位及不同發育時期的水稻組織，序列長度介於 48-1,733 bp 之間，提供來源者包括中國大陸、美國、韓國及印度等不同國家。

EST 的聚集與接合

由於不同的基因所表現的 mRNA 量不同，選殖時表現量多的基因較容易被選殖出來且產生較多數 EST 序列，表現量少的基因則較不易獲得 EST 序列。然而由於 EST 序列是隨機取樣之各轉錄

表 1. NCBI 之 dbEST 資料庫中受稻熱病菌誘導之水稻 EST 的資訊摘要 (至 2007/7/31 止)

Table 1. Information summary of rice ESTs induced by *Magnaporthe grisea* from NCBI-dbEST database (update time: July 31, 2007)

| Genbank Acc. | File name | EST no. | Entry creaty | Length (bp) | Variety ^z | Tissue | Developmental stage | Country |
|--------------------------------------|-----------|---------|------------------------|-------------|---|----------------------|--------------------------|---------|
| BF889457-BF889486 | BF-1 | 14 | 2001/5/1 | 173-577 | Drew (japonica) | leaf | 16-day-old seedling | USA |
| BI118615-BI118864 | BI-2 | 250 | 2001/6/26 | 152-1175 | Yuanfeengzao (indica) | leaf | 3 week seedling | USA |
| BI807369-BI811654, BI811655-BI813841 | BI-1 | 6437 | 2001/10/2 2001/11/1 | 146-860 | <i>Oryza sativa</i> | leaf | mature stage | China |
| BM037978-BM039065 | BM-2 | 1088 | 2001/11/6 | 60-1086 | <i>O. sativa</i> | leaf | mature stage | China |
| BM418528-BM422311 | BM-1 | 3784 | 2002/1/28 | 169-1045 | <i>O. sativa</i> | leaf | mature stage | China |
| AF521978-AF522023 | AF | 46 | 2002/7/30 | 60-600 | Bianyiyuexiangzhan (indica) | leaf | seedling five leaf stage | China |
| BQ905903-BQ909262 | BQ | 3360 | 2002/8/19 | 58-800 | <i>O. sativa</i> | leaf | mature stage | China |
| BU666934-BU667367 | BU | 434 | 2002/9/30 | 48-788 | <i>O. sativa</i> | leaf | mature stage | China |
| CB617709-CB635999 | CBir36 | 18241 | 2003/4/8 | 159-979 | IR36 (japonica) | leaf | 3 week seedling | USA |
| CB636000-CB686047 | CBbare | 49403 | 2003/4/8-9 | 152-973 | Nipponbare (japonica) | leaf | 3 week seedling | USA |
| BF145163-BF145216, BF108309-BF108365 | BF-2 | 111 | 2003/12/30 | 115-798 | Minghui 63 (indica) | 14 different tissues | | China |
| CD347671-CD347779 | CD | 109 | 2004/3/5 | 62-815 | <i>Oryza minuta</i> 101144 (wild) | leaf | Vegetative stage | Korea |
| CX727819-CX728959 | CX | 1141 | 2005/1/21 | 152-686 | Nipponbare (japonica) IR 36 (japonica) | leaf | 3 week seedling | USA |
| DN475717-DN475431 | DN | 287 | 2005/3/9 | 143-1733 | <i>O. s ativa</i> (indica cultivar-grop) | 13 different tissues | | Indica |
| Total | | 84705 | | 48-1733 | | | | |

^z *O. sativa*: variety name wasn't described by submitter.

產物 mRNA，表現量多的 EST 定序可能具有較高的重複性而導致產生多餘序列，故可能會有多個 EST 代表同一個轉錄產物。因此在進行 EST 分析之前，為維持序列的品質，通常都會運用聚集與接合的方法，減少 EST 的重疊性，並將其連接成更長、更高質量的序列。這種分析之前的常規化處理對 EST 功能的識別、剪接產物的區分和基因表達圖譜的分析都有很大的幫助，進而有利於功能註解之進行 (Adams *et al.* 1995)。聚集之目的就是將來自於同一個基因之具有重疊部分的 EST 歸至單一的群 (cluster)，而無法歸群者則單獨成群，如此可建立一個非多餘性的序列資料庫 (Kohchi *et al.* 1995; Bonaldo *et al.* 1996)；然後透過接合，即將同一群內的多個序列，經由剪切 (clip) 5' 和 3' 端之低品質區域並計算讀取間的重疊性，再依重疊計分由大到小依序將讀取序列連結成較長的接合序列。EST 經過聚集與接合所得的單獨成群之序列和接合序列，都屬於彼此之間沒有重疊的單一序列。本研究使用 CAP3 序列接合程式 (Huang & Madan 1999)，對受稻熱病菌誘導之水稻 EST 序列，分別以分開不同 cDNA 基因庫來進行聚集及接合，然後再以所得的單一序列，進行後續的基因功能比較分析。

EST 單一序列之功能的註解、分類及表現量統計

Chiu *et al.* (2007) 以受褐飛蝨誘導之水稻 EST 資料為材料，成功建立了一套完整的 EST 功能分析流程，可快速進行 EST 功能的註解、分類及表現量統計：(1) 首先去比對 TIGR Gene Index (TGI) 資料庫 (<http://compbio.dfci.harvard.edu/tgi>)，以進行功能註解；(2) 藉由搜尋結果連結至相對應之 GO (<http://www.geneontology.org/>) 分類代碼及敘述，可進行功能註解之粗分類；(3) 再經由 GO 分類代碼連結至相對應之 MIPS (Munich Information Center for Protein Sequences; <http://mips.gsf.de/>) 分類代碼目錄，以進行功能註解之細分類；(4) 最後計算各類功能類別的個數統計及百分比計算。TGI 蒐集的基因資料不僅是簡單地對 NCBI 之 GenBank 資料庫的基因和 EST 數據按物種分類，也提取 GenBank 中不同物種的 EST 和註解的基因序列，並通過聚集、接合等分析過程產生單一的、具有高可信度的假設性一致序列 (tentative consensus sequence, TC)，此 TC 可用來提供推測基因之功能註解。GO 主要提供結構及動態模式的控制字彙和分類系統，真核生物的基因和蛋白質也可透過 GO 的整合性分類系統得到分子功能 (molecular function)、生物程序 (biological process)、細胞成分 (cellular component) 三個層次的註解，可對基因組資訊進行簡單分類及描述 (GO 2006)。MIPS 分類系統的功能分類目錄 (MIPS_FunCat; http://mips.gsf.de/proj/funcatDB/search_main_frame.html) 由 22 個主要功能種類所組成，可用來對不同生物之基因組進行註解之分類，是研究者對基因及其產物進行功能分類時常參考的資料庫 (Andreas *et al.* 2004; Mewes *et al.* 2006)。因此，本研究以 Chiu *et al.* (2007) 所建立之 EST 功能分析流程為基礎，將受稻熱病菌誘導之水稻 EST 單一序列的各基因庫 FASTA 格式檔案，先去比對 TGI 中 Rice (*O. sativa*) 之 TC 資料庫 (搜尋比對之 E 值內設為 10)，由比對結果找出最低 E 值的 TC 以進行功能註解，再根據 GO 之三個層次分類及 MIPS_FunCat 之 22 個主要功能種類目錄，將比對到的結果，進一步進行基因各類功能表現量的統計分析。由於這部份的 EST 資料量非常龐大 (共 84,705 筆)，為減少 EST 功能註解所需的人力及時間，本研究同時利用 Perl 語言撰寫程式，將整個流程建立了自動化分析系統 (Lu *et al.* 2006)，以快速準確地完成 EST 全程的功能分析工作。

利用多變數統計技術進行基因庫與功能間關係之探討

由於受稻熱病菌誘導之水稻 EST 資料涉及到多個 cDNA 基因庫 (內含不同品種、不同植株部位、不同發育時期之水稻 EST 資料)，為探討各功能類別與基因庫間之關係，故在同時考慮多個 EST 功能類別下將所有 cDNA 基因庫依其相似或相異程度，首先利用集群分析的全連法 (complete

linkage) 進行分群，以得到 cDNA 基因庫的樹狀圖 (dendrogram)。集群分析是一種一般邏輯程序，能根據相似性與相異性，客觀地將相似者歸在同一群內，其一般原則是使同一集群內的個體差異性最小，而不同集群間的個體則具有最大差異性；而集群分析中可選擇多種演算法，其中全連法是以最遠的點之間的距離做為集群間的距離，又稱為遠鄰法 (farthest neighbor)，以最遠的個體來判斷相似性，最不容易造成連鎖，故形成的樹最長 (故又稱長臂法)，可使個體之間形成最緊湊的群，在生物學上最好解釋 (Lepš & Šmilauer 1999)。

接著本研究也進行對應分析，以得到 cDNA 基因庫與功能類別在二維空間的圖示關係。對應分析為 Benzécri 於 1973 年所提出，又稱為相互平均法 (reciprocal averaging; Hill 1973)，用於分析簡單二維及多維表格的行與列之間的對應關係，是一種廣泛運用到生態學 (尤其是植物生態學) 的多變數技術。該法現已大量應用到其他各方面研究領域 (Beh 2004)，包括基因組分析上之各種應用 (Gupta & Ghosh 2001; Tekaiia *et al.* 2002; Tan *et al.* 2004)。對應分析對於數據的要求比較簡單，不需要有嚴格的線性關係，也無須限定所處理的資料是連續或離散變數，只需要是一個列聯表的資料形式，即可將多維之變數或樣本於一個低維空間上以一個序列分佈圖 (ordination diagram) 呈現，以研究出其變數間或樣本間之關係，也可將變數和樣本放在同一張圖 (稱為雙標圖, biplot) 上來同時表達出變數與樣本二者間相互對應的關係 (Lepš & Šmilauer 1999; Oksanen 2004)。本研究藉由對應分析所得之 cDNA 基因庫的序列分佈圖，利用各基因庫點落在圖上之距離遠近，來輔助上述集群分析之樹狀圖的分群結果判斷，倘發現點間有群團現象表示可能是同一群；同時進而繪製 cDNA 基因庫與功能類別在同一張圖上的雙標圖，以辨識出 cDNA 基因庫分群結果與功能類別兩者之間的相互關係，倘某功能類別與該基因庫群的關係密切但與其他群關係較遠，表示其為該基因庫群之所以異於其他群的主要功能類別。

本研究的集群分析和對應分析都是使用 STATISTICA 統計軟體 (Statsoft Inc. 2002)，在 Window 作業系統個人電腦上執行。

結 果

EST 的聚集與接合

擷取自 NCBI 之 dbEST 資料庫中受稻熱病菌誘導之水稻 EST 序列共 84,705 條，分屬於 14 個 cDNA 基因庫，分別利用 CAP3 序列接合程式進行聚集及接合後，各基因庫分別獲得 14-14,281 條單一序列 (表 2)。當基因庫內 EST 序列條數較多時，序列之間的覆蓋率相對較高，經 CAP3 序列接合程式處理後所減少之序列多餘性的量亦相對較多，如 CBir36 基因庫內含 18,241 條 EST 序列，所減少之多餘性可達 61% (即所產生的單一序列所佔比例為 39%)，而 CBbare 基因庫的 EST 資料量近 5 萬筆，其減少之多餘性更高達 71% (單一序列所佔比例為 29%)。

EST 單一序列之功能的註解、分類及表現量統計

將受稻熱病菌誘導之水稻 EST 經聚集及接合後所產生的單一序列，藉由比對 TC 所搜尋之結果進行功能註解，再連結對應到 GO 功能分類系統，並計算 GO 分類系統之功能類別基因表現的個數及百分比，結果如表 3。結果顯示來自不同品種、不同植株部位及不同發育時期受稻熱病菌誘導之水稻 EST，其所屬的功能分類在 GO 粗分類層次上之表現差異不大，整體來說皆以生物程序 (biological process) 功能為主 (34.2-43.1%)，其次是細胞組成 (cellular component) 功能 (29.3-36.8%)，而分子功能 (molecular function) 相關基因則最少 (25.0-32.2%)。

進而利用 GO 分類代碼與 MIPS 分類目錄代碼間的對應關係，得到在 MIPS 功能分類系統下各 cDNA 基因庫之更細分的功能類別所佔 EST 個數及其比例，示如表 4。根據這些 cDNA 基因庫的來源品種（表 1），再區分秈稻 (indica rice)、粳稻 (japonica rice) 及野生稻 (wild rice) 來比較其間差異（圖 1），發現各品種受稻熱病菌誘導之水稻 EST，都具有代謝 (metabolism)、能量 (energy)、細胞週期及 DNA 修飾 (cell cycle and DNA processing)、蛋白質循環轉變（即蛋白質的折疊、修飾和傳送終點之生命週期，protein fate / folding, modification, destination）、細胞生長和形態發生 (cell growth / morphogenesis)、次細胞定位 (subcellular localization)、蛋白質結合 (protein binding)、運輸促進作用 (transport facilitation) 和無法功能歸類之蛋白質 (unclassified proteins) 等功能。除無法功能歸類之蛋白質外，各水稻品種均以次細胞定位之功能的基因表現量最高，但以野生稻所佔比例 (27.5%) 高於秈稻 (21.8%) 和粳稻 (18.9%)。除此，野生稻的代謝 (7.3%)、細胞生長和形態發生 (2.5%) 及運輸促進作用 (10.0%) 之功能的基因量，也高於秈稻（各 5.9、0.1、5.7%）和粳稻（各 6.1、0.2、3.6%）。然而，野生稻的能量 (3.0%)、蛋白質循環轉變 (0.5%) 及蛋白質結合 (1.5%) 之功能的基因表現量，則低於秈稻（各 7.0、7.2、5.7%）和粳稻（各 4.1、7.1、3.6%），尤其是蛋白質循環轉變功能之基因的表現差異最為明顯。另外，秈稻的能量功能之基因表現量略高於粳稻，而粳稻的蛋白質合成 (protein synthesis) 功能之基因表現量 (4.1%) 略高於秈稻 (1.1%)。各品種無法功能歸類之蛋白質所佔比例高達 36.7-45.5%，表示仍存在有未知功能之基因，故值得未來再進一步探討這些無法歸類蛋白質之功能為何。至於圖 1 中的未知水稻品種 EST，因序列提供者並未說明品種為何，但從本研究中並未發現它和其他已知品種之功能基因表現有特別明顯相似或相異，故難以將它與已知品種來相互作比較，不擬作特別探討。

表 2. 受稻熱病菌誘導之水稻 EST 序列經 CAP3 聚集及接合的結果

Table 2. Clustering and assembly results with CAP3 software in rice ESTs induced by *Magnaporthe grisea*

| cDNA library | EST no. | EST length (bp) | Contig no. | Contig length (bp) | Singleton no. | Singleton (%) | Singleton length (bp) | Unique ^z sequence no. | Unique ^z sequence % |
|--------------|---------|-----------------|------------|--------------------|---------------|---------------|-----------------------|----------------------------------|--------------------------------|
| BF-1 | 14 | 173-577 | 0 | 0 | 14 | 100 | 173-577 | 14 | 100 |
| AF | 46 | 60-600 | 2 | 253-337 | 40 | 87 | 60-600 | 42 | 91 |
| CD | 109 | 62-815 | 13 | 248-570 | 69 | 63 | 104-815 | 82 | 75 |
| BF-2 | 111 | 115-798 | 1 | 671 | 109 | 98 | 115-798 | 110 | 99 |
| BI-2 | 250 | 152-1175 | 38 | 285-1052 | 166 | 66 | 152-1175 | 204 | 82 |
| BU | 434 | 48-820 | 32 | 176-827 | 333 | 77 | 48-794 | 365 | 84 |
| BM-2 | 1088 | 60-1086 | 122 | 234-1083 | 707 | 65 | 60-821 | 829 | 76 |
| BQ | 3360 | 58-958 | 457 | 25-1088 | 1707 | 51 | 73-958 | 2164 | 64 |
| BM-1 | 3784 | 132-1070 | 530 | 316-1023 | 1798 | 48 | 144-1070 | 2328 | 62 |
| BI-1 | 6437 | 146-860 | 994 | 241-1085 | 2642 | 41 | 146-860 | 3636 | 56 |
| CBir36 | 18241 | 159-979 | 2873 | 208-4479 | 4257 | 23 | 166-952 | 7130 | 39 |
| CBbare | 49403 | 152-973 | 6987 | 167-4379 | 7294 | 15 | 152-952 | 14281 | 29 |
| DN | 287 | 143-1733 | 9 | 297-1411 | 180 | 63 | 143-1733 | 189 | 66 |
| CX | 1141 | 152-686 | 148 | 242-925 | 643 | 56 | 152-686 | 791 | 69 |
| Total | 84705 | | 12206 | | 19959 | | | 32165 | |

^z Unique sequences = contigs and singletons.

表 3. 在 GO 功能分類系統下受稻熱病菌誘導之水稻 EST 的功能類別統計

Table 3. Statistics of functional classification for rice ESTs induced by *Magnaporthe grisea*, based on GO functional category

| cDNA library | GO functional category | | | | | |
|--------------|------------------------|------|--------------------|------|--------------------|------|
| | Molecular function | | Biological process | | Cellular component | |
| | EST no. | % | EST no. | % | EST no. | % |
| BF-1 | 11 | 28.9 | 13 | 34.2 | 14 | 36.8 |
| AF | 17 | 31.5 | 19 | 35.2 | 18 | 33.3 |
| BF-2 | 77 | 27.1 | 108 | 38.0 | 99 | 34.9 |
| BI-1 | 1884 | 30.5 | 2386 | 38.6 | 1904 | 30.8 |
| BI-2 | 174 | 30.9 | 213 | 37.8 | 177 | 31.4 |
| BM-1 | 1345 | 29.8 | 1754 | 38.8 | 1420 | 31.4 |
| BM-2 | 492 | 28.5 | 671 | 38.8 | 566 | 32.7 |
| BQ | 1168 | 30.0 | 1484 | 38.1 | 1246 | 32.0 |
| BU | 128 | 27.6 | 185 | 39.9 | 151 | 32.5 |
| CBbare | 9021 | 32.2 | 10681 | 38.1 | 8334 | 29.7 |
| CBir36 | 4761 | 31.9 | 5801 | 38.8 | 4372 | 29.3 |
| CD | 29 | 27.9 | 38 | 36.5 | 37 | 35.6 |
| CX | 554 | 27.4 | 744 | 36.8 | 722 | 35.7 |
| DN | 40 | 25.0 | 69 | 43.1 | 51 | 31.9 |

將受稻熱病菌誘導之水稻 EST，再根據各 cDNA 基因庫的來源組織和感染時期（表 1），區分為幼苗期葉片（AF 基因庫）、分蘖期葉片（如 BF-1、BI-2、CBir36、CBbare、CD 及 CX 基因庫，含發芽後 16 天、3 週及營養生長期的葉片）、抽穗成熟期葉片（如 BI-1、BM-1、BM-2、BQ 及 BU 基因庫）以及不同部位組織（如 BF-2 及 DN 基因庫，其發育時期未知），以檢視其間的功能基因表現差異，結果如圖 2 所示。結果顯示，隨著植株成長，水稻葉片愈近生長後期，受稻熱病菌侵染後，代謝和能量之相關功能的基因表現量愈高，但次細胞定位功能之基因的表現量似乎有減少現象；幼苗期的葉片甚至完全沒有表現出能量功能之基因量，但次細胞定位功能之基因表現量卻高於生長後期的葉片，而且蛋白質循環轉變功能之基因表現量居次，也明顯高於生長後期葉片之表現。相對於固定葉片組織，混合了不同組織的水稻 EST 似乎表現出較低的代謝功能（3.4%）與較高的能量功能（10.7%），而其他功能之基因表現量則相異不大。

利用多變數統計技術進行基因庫與功能間關係之探討

進而針對 MIPS 功能分類系統之分析結果，分別以集群分析和對應分析對 14 個 cDNA 基因庫結果進行相異性分群。根據所得的樹狀圖（圖 3-A）以歐氏距離（Euclidean distances）約 22.4 之距離位點劃分和序列分布圖（圖 3-B）之綜合結果來看，大致可將 14 個 cDNA 基因庫依其相異性分成三群（I、II 和 III）；然而，發現第 I 群 cDNA 基因庫之分布過於緊密，難以看出此群 cDNA 基因庫間之差異性，於是重新檢視 EST 資料，發現可能是由於 AF 與 BF-1 基因庫之單一序列條數過少，各 42 及 14 條（表 2），恐怕會影響了整體分析之準確度，故本研究進而再將 AF 與 BF-1 這二個基因庫移除後，重新分析。

表 4. 在 MIPS 功能分類系統下受稻熱病菌誘導之水稻 EST 的功能類別統計

Table 4. Statistics of functional classification for rice ESTs induced by *Magnaporthe grisea*, based on MIPS functional category

| cDNA library | MIPS functional category ^z | | | | | | | | | | | | | | | | | |
|--------------|---------------------------------------|------|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|------|------|-----|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 11 | 13 | 14 | 25 | 30 | 40 | 63 | 65 | 67 | 99 |
| | — EST no. — | | | | | | | | | | | | | | | | | |
| AF | 3 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 9 | 2 | 0 | 3 | 10 |
| BF-1 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 3 | 1 | 1 | 0 | 1 | 0 | 6 | 2 | 0 | 0 | 4 |
| BF-2 | 6 | 17 | 5 | 0 | 7 | 37 | 7 | 4 | 16 | 0 | 0 | 1 | 9 | 91 | 30 | 0 | 7 | 46 |
| BI-1 | 253 | 210 | 60 | 19 | 106 | 269 | 97 | 67 | 77 | 14 | 36 | 16 | 79 | 805 | 269 | 0 | 156 | 903 |
| BI-2 | 39 | 34 | 19 | 1 | 8 | 12 | 5 | 16 | 7 | 1 | 9 | 3 | 6 | 106 | 32 | 0 | 13 | 101 |
| BM-1 | 204 | 168 | 58 | 20 | 67 | 179 | 63 | 43 | 51 | 13 | 18 | 13 | 61 | 610 | 185 | 0 | 126 | 686 |
| BM-2 | 88 | 82 | 18 | 2 | 29 | 73 | 29 | 16 | 24 | 5 | 11 | 7 | 29 | 241 | 82 | 0 | 47 | 226 |
| BQ | 161 | 138 | 50 | 12 | 87 | 184 | 74 | 53 | 58 | 12 | 22 | 8 | 60 | 635 | 190 | 0 | 105 | 553 |
| BU | 34 | 27 | 17 | 1 | 13 | 26 | 16 | 3 | 17 | 0 | 12 | 3 | 1 | 93 | 29 | 0 | 7 | 74 |
| CBbare | 1221 | 780 | 567 | 135 | 400 | 1447 | 678 | 393 | 306 | 144 | 139 | 65 | 565 | 4267 | 1688 | 11 | 822 | 5623 |
| CBir36 | 705 | 445 | 345 | 63 | 222 | 902 | 356 | 235 | 187 | 61 | 81 | 30 | 282 | 2272 | 1058 | 3 | 474 | 2622 |
| CD | 6 | 2 | 4 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 20 | 2 | 0 | 3 | 22 |
| CX | 74 | 109 | 20 | 8 | 56 | 95 | 49 | 24 | 24 | 2 | 4 | 3 | 48 | 416 | 83 | 0 | 47 | 269 |
| DN | 1 | 5 | 1 | 0 | 1 | 3 | 2 | 0 | 1 | 0 | 0 | 0 | 1 | 13 | 5 | 0 | 5 | 19 |
| | — % — | | | | | | | | | | | | | | | | | |
| AF | 5.4 | 0.0 | 5.0 | 0.0 | 0.0 | 12.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | 26.1 | 5.0 | 0.0 | 6.1 | 37.5 |
| BF-1 | 0.0 | 0.0 | 0.0 | 0.0 | 7.1 | 4.8 | 0.0 | 7.1 | 2.4 | 2.4 | 0.0 | 7.1 | 0.0 | 23.8 | 9.5 | 0.0 | 0.0 | 35.7 |
| BF-2 | 4.5 | 9.4 | 2.9 | 0.0 | 1.3 | 9.2 | 1.8 | 0.2 | 1.2 | 0.0 | 0.0 | 0.8 | 2.6 | 21.9 | 3.7 | 0.0 | 4.7 | 35.7 |
| BI-1 | 9.0 | 5.9 | 0.9 | 0.8 | 2.0 | 8.4 | 1.6 | 1.2 | 1.1 | 0.2 | 0.3 | 0.7 | 1.5 | 17.1 | 6.5 | 0.0 | 5.4 | 37.4 |
| BI-2 | 11.2 | 6.7 | 1.8 | 0.2 | 2.5 | 3.5 | 1.5 | 3.3 | 2.9 | 1.0 | 0.2 | 0.2 | 0.2 | 17.5 | 5.8 | 0.0 | 4.4 | 37.2 |
| BM-1 | 10.2 | 5.9 | 1.1 | 0.9 | 1.6 | 7.5 | 2.0 | 1.1 | 0.9 | 0.5 | 0.2 | 0.5 | 1.7 | 17.6 | 5.7 | 0.0 | 6.3 | 36.2 |
| BM-2 | 11.6 | 8.7 | 1.3 | 0.2 | 2.1 | 7.4 | 2.7 | 1.0 | 1.5 | 0.5 | 0.3 | 0.5 | 2.3 | 18.4 | 6.7 | 0.0 | 4.8 | 30.0 |
| BQ | 9.3 | 5.7 | 1.0 | 0.7 | 2.4 | 6.7 | 2.3 | 1.4 | 1.0 | 0.4 | 0.2 | 0.4 | 1.8 | 19.6 | 5.4 | 0.0 | 5.7 | 36.0 |
| BU | 14.3 | 12.7 | 0.5 | 0.1 | 1.2 | 7.6 | 0.3 | 1.3 | 1.5 | 0.0 | 0.4 | 0.2 | 0.2 | 16.4 | 4.5 | 0.0 | 1.5 | 37.2 |
| CBbare | 8.4 | 3.8 | 1.2 | 0.7 | 1.8 | 7.6 | 1.5 | 1.3 | 0.8 | 0.3 | 0.3 | 0.4 | 1.5 | 15.1 | 7.0 | 0.1 | 5.3 | 42.9 |
| CBir36 | 9.1 | 3.7 | 1.5 | 0.6 | 2.1 | 9.6 | 1.6 | 1.2 | 0.9 | 0.3 | 0.3 | 0.4 | 1.6 | 15.2 | 8.4 | 0.0 | 5.7 | 37.8 |
| CD | 7.3 | 3.0 | 2.4 | 0.0 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 27.5 | 1.5 | 0.0 | 10.0 | 45.5 |
| CX | 7.0 | 9.1 | 1.1 | 0.9 | 5.5 | 6.5 | 2.1 | 0.7 | 0.9 | 0.1 | 0.1 | 0.3 | 2.1 | 21.6 | 4.9 | 0.0 | 3.5 | 33.3 |
| DN | 2.4 | 11.9 | 2.4 | 0.0 | 0.5 | 3.4 | 3.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 1.6 | 21.9 | 8.1 | 0.0 | 7.6 | 36.4 |

^z Functional category : 1- metabolism, 2- energy, 3- cell cycle and DNA processing, 4- transcription, 5- protein synthesis, 6- protein fate (folding, modification, destination), 8- cellular transport and transport mechanisms, 10- cellular communication / signal transduction mechanism, 11- cell rescue, defense and virulence, 13- ionic homeostasis, 14- cell growth / morphogenesis, 25- development (systemic), 30- control of cellular organization, 40- subcellular localization, 63- protein binding, 65- storage protein, 67- transport facilitation, 99- unclassified proteins.

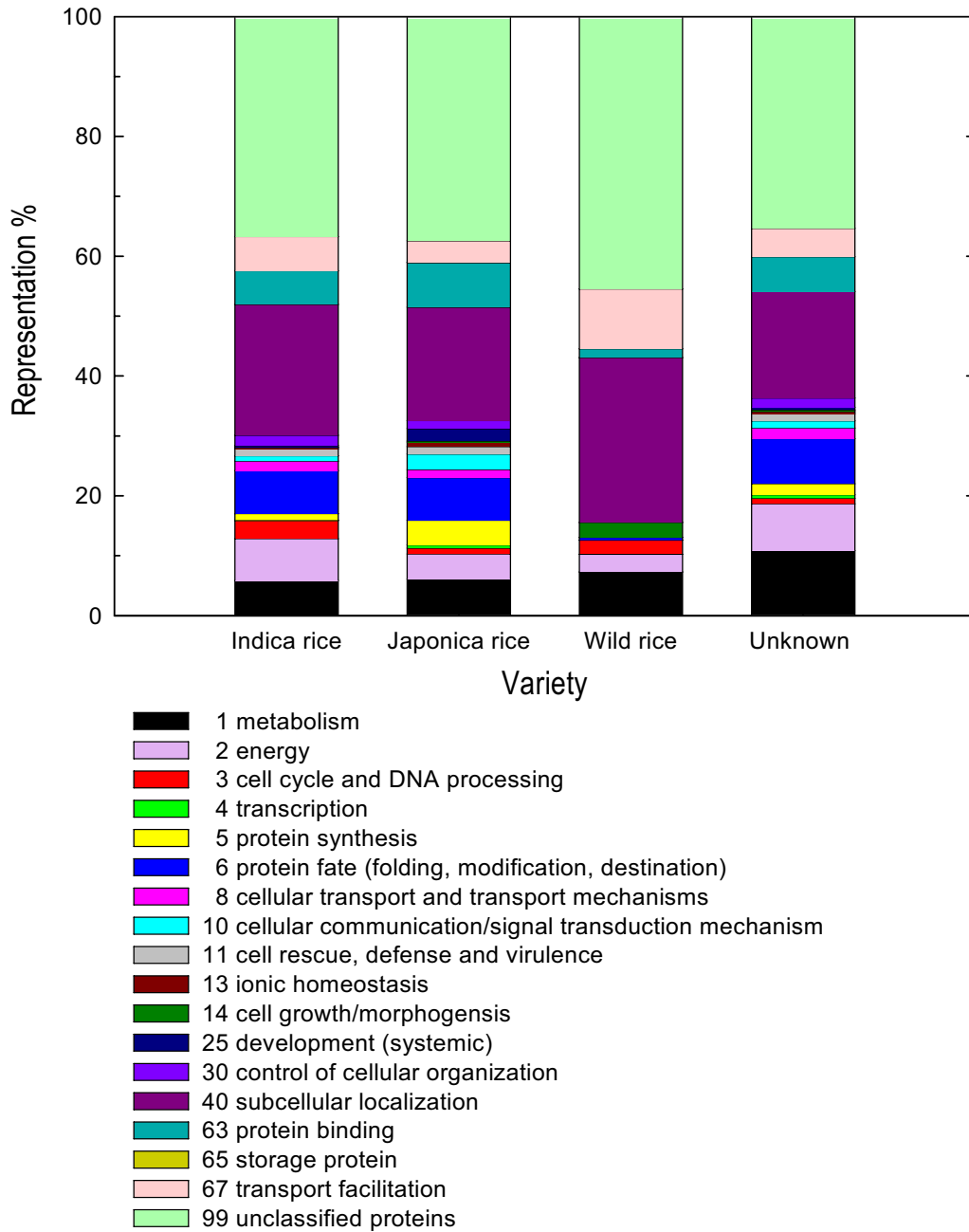


圖 1. 在 MIPS 功能分類系統下受稻熱病菌誘導之不同水稻品種之 EST 的比較。

Fig. 1. Comparison of rice EST induced by *Magnaporthe grisea* among different rice varieties, based on MIPS functional category system.

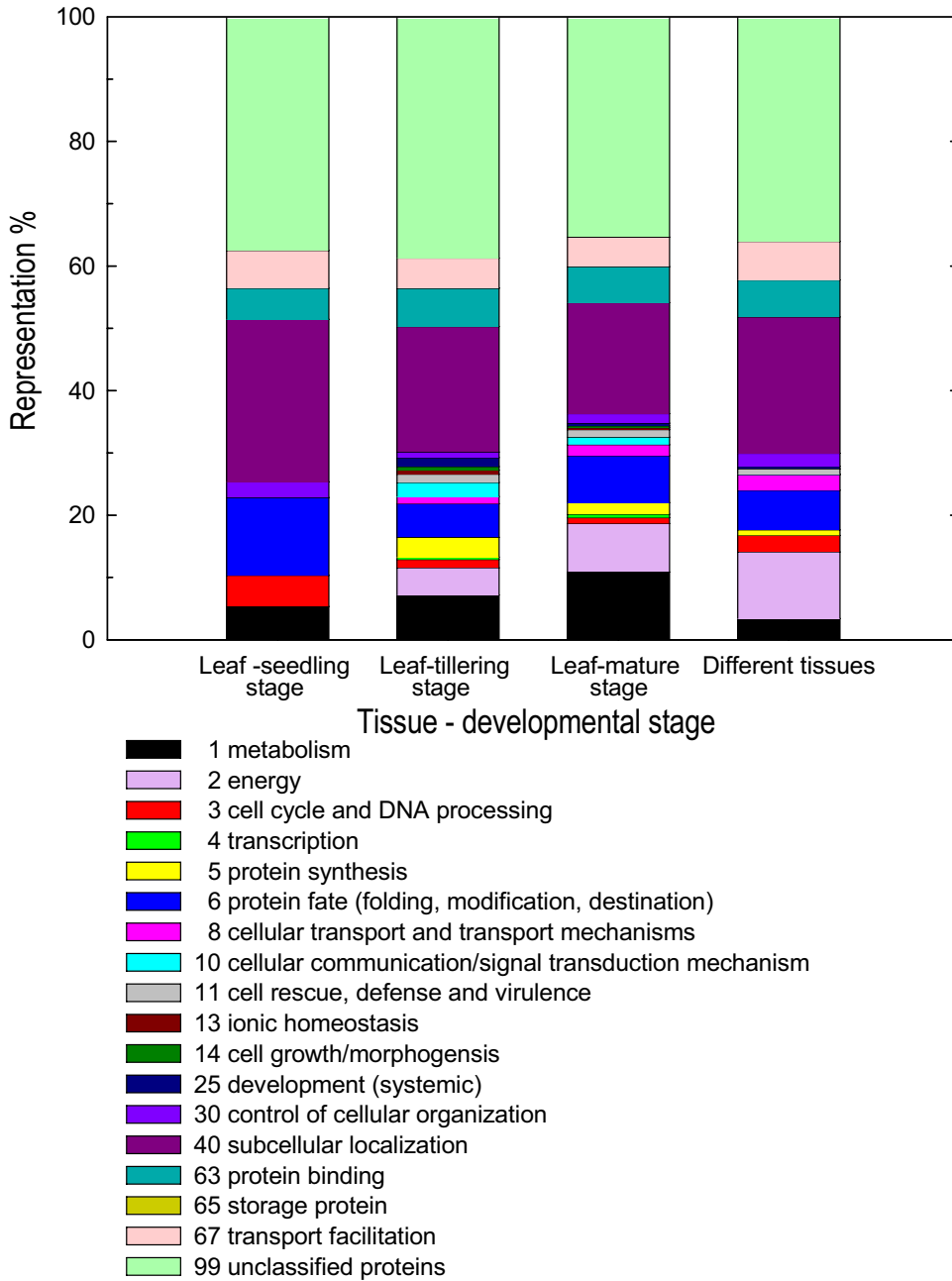


圖 2. 在 MIPS 功能分類系統下受稻熱病菌誘導之不同水稻植株部位與不同發育時期之 EST 的比較。
 Fig. 2. Comparison of rice EST induced by *Magnaporthe grisea* among different rice tissues and developmental stages, based on MIPS functional category system.

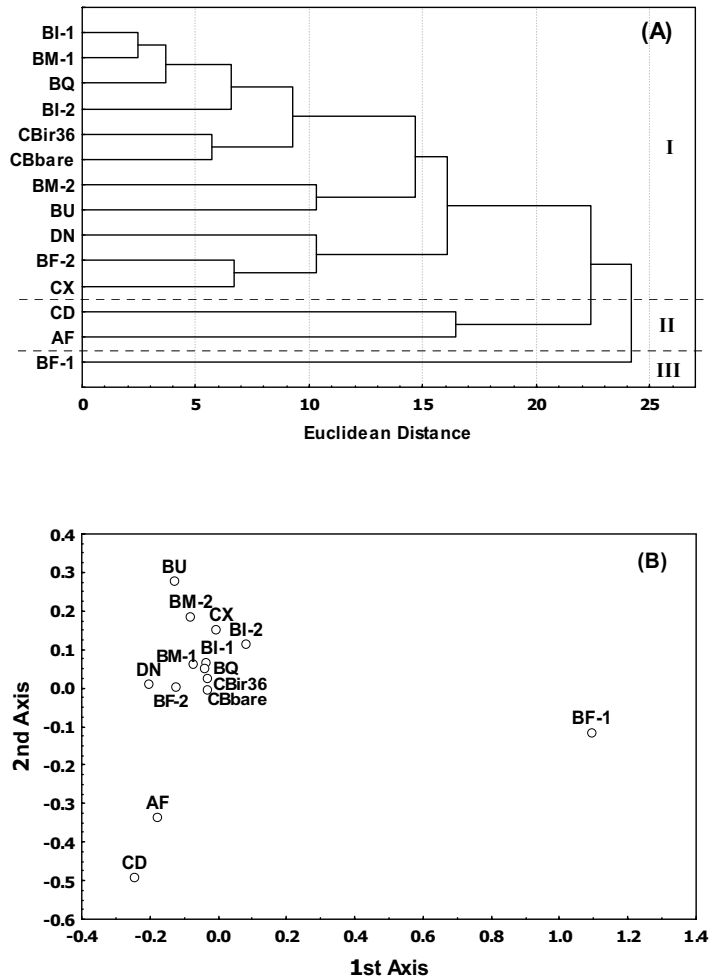


圖 3. 受稻熱病菌誘導之 14 個水稻 cDNA 基因庫的基因表現之分群結果。(A) 集群分析採全連法之樹狀圖，(B) 對應分析之序列分布圖。

Fig. 3. Grouping results of gene expression of 14 rice cDNA libraries induced by *Magnaporthe grisea*. (A) Dendrogram of clustering analysis using complete linkage, (B) Ordination diagram of correspondence analysis.

移除 AF 與 BF-1 基因庫後，在同時考慮多個 EST 功能類別下，12 個 cDNA 基因庫所得的樹狀圖（圖 4-A）以歐氏距離約 16.1 之距離位點劃分和序列分布圖（圖 4-B）之綜合結果來看，大致可將 12 個 cDNA 基因庫依其相異性分成 3 群（以 I、II 和 III 表示）：第 I 群包含 BI-1, BM-1, BQ, BI-2, CBir36, CBbare, BM-2, BU 此八個 cDNA 基因庫，除了 BI-1, BM-1, BM-2, BQ 及 BU 為未知品種外，其他序列來源均屬於單一品種、單一感染部位；第 II 群包含 DN, BF-2, CX 此三個 cDNA 基因庫，序列來源分別混合不同品種之不同組織、混合不同品種之葉片組織；第 III 群則僅有 CD 一個 cDNA 基因庫，屬於野生稻之葉片組織。

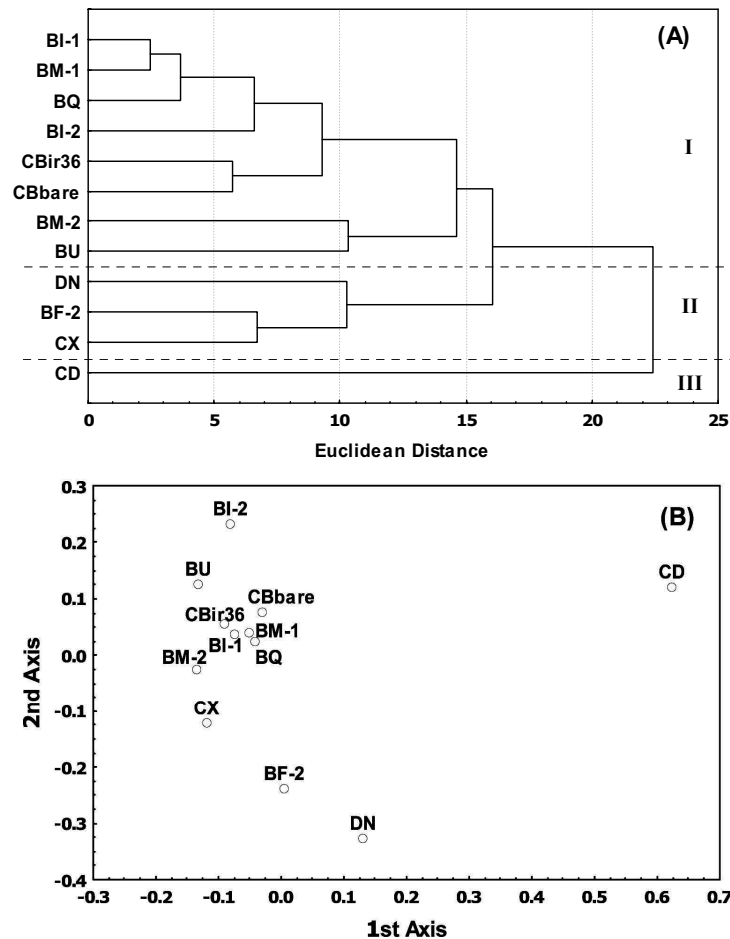


圖 4. 受稻熱病菌誘導之 12 個水稻 cDNA 基因庫的基因表現之分群結果。(A) 集群分析採全連法之樹狀圖，(B) 對應分析之序列分布圖。

Fig. 4. Grouping results of gene expression of 12 rice cDNA libraries induced by *Magnaporthe grisea*. (A) Dendrogram of clustering analysis using complete linkage, (B) Ordination diagram of correspondence analysis.

利用對應分析再將 12 個 cDNA 基因庫與 EST 功能類別放在同一張圖上所得之雙標圖如圖 5 所示，可看出基因庫分群結果與其功能類別間之相互關係：由於細胞間聯絡及訊息傳導之機制 (cellular communication / signal transduction mechanism)、離子恆定性 (ionic homeostasis) 與貯藏性蛋白 (storage protein) 之功能與第 I 群 cDNA 基因庫之關係較近，但與其他兩群之 cDNA 基因庫關係較遠，因此推估第 I 群 cDNA 基因庫之 EST 可能主要是受到大量表現出這 3 種蛋白質功能，而導致與其他兩群的 cDNA 基因庫有明顯差異；而第 II 群 cDNA 基因庫之 EST 可能主要是受到大量表現出能量、細胞間聯絡及訊息傳導之機制與細胞組織之控制 (control of cellular organization) 之功能，而異於其他兩群；至於第 III 群 cDNA 基因庫之 EST 則可能主要是受到表現出大量的細胞生長／形態發生和傳輸促進功能，而異於其他兩群，尤其以細胞生長和形態發生之功能的影響最大。若檢視各群之 EST 以 MIPS 功能分類系統下各功能類別所佔比例之平均值 (表 5)，可看出第 I 群之

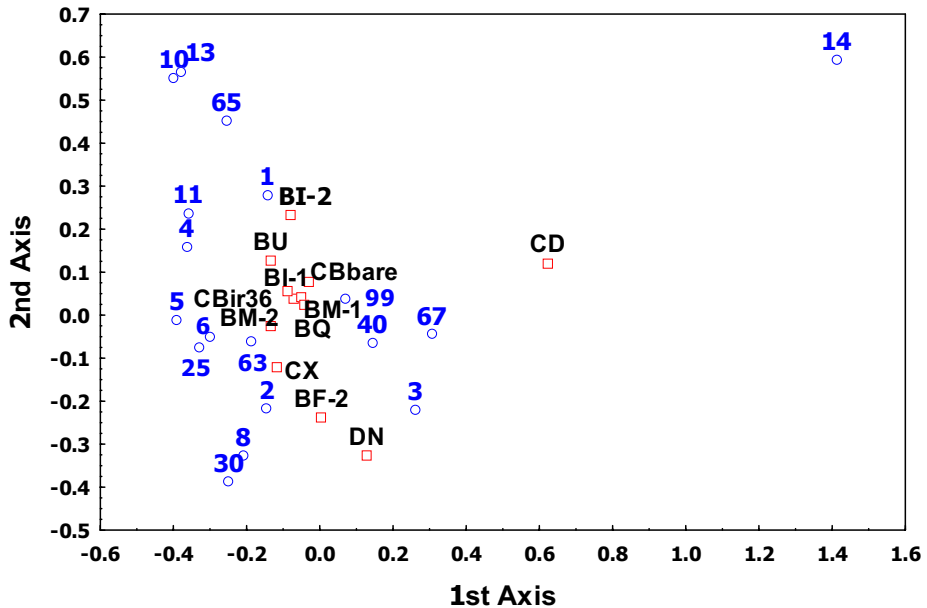


圖 5. 利用對應分析顯示 12 個水稻 cDNA 基因庫與其 MIPS 分類系統所得之功能類別之關係的雙標圖。(cDNA 基因庫：BI-1, BM-1, BQ, BI-2, CBir36, CBbare, BM-2, BU, DN, BF-2, CX, CD；功能類別：1-代謝, 2-能量, 3-細胞週期及 DNA 修飾, 4-轉錄, 5-蛋白質合成, 6-蛋白質命運/折疊, 修飾, 傳送終點, 8-細胞傳輸及其機制, 10-細胞間聯絡及訊息傳導之機制, 11-細胞援救、防禦和毒性, 13-離子恆定性, 14-細胞生長/形態發生, 25-發育(系統性的), 30-細胞組織之控制, 40-次細胞定位, 63-蛋白質結合, 65-貯藏性蛋白, 67-傳輸促進作用, 99-功能未歸類之蛋白質)

Fig. 5. A biplot showing the interrelationships between 12 rice cDNA libraries and their functional classifications using correspondence analysis based on MIPS functional category system. (cDNA library:BI-1, BM-1, BQ, BI-2, CBir36, CBbare, BM-2, BU, DN, BF-2, CX, CD; Functional category:1- metabolism, 2- energy, 3- cell cycle and DNA processing, 4- transcription, 5- protein synthesis, 6- protein fate / folding, modification, destination, 8- cellular transport and transport mechanisms, 10- cellular communication / signal transduction mechanism, 11- cell rescue, defense and virulence, 13- ionic homeostasis, 14- cell growth / morphogenesis, 25- development (systemic), 30- control of cellular organization, 40- subcellular localization, 63- protein binding, 65- storage protein, 67- transport facilitation, 99- unclassified proteins)

EST 的貯藏性蛋白功能之相關基因雖然表現量不高 (僅 0.02%)，但於其他兩群之表現量均為 0%；而細胞間聯絡及訊息傳導之機制功能之相關基因表現量 (1.48%) 高於其他兩群 (0-0.30%)，以及離子恆定性功能 (0.40%) 高於其他兩群 (0-0.03%)；第 II 群之 EST 之能量功能 (10.14%) 高於其他兩群 (2.95-6.63%)，而細胞間聯絡及訊息傳導之機制功能 (2.38%) 亦高於其他兩群 (0-1.68%)，以及細胞組織之控制 (2.11%) 也是高於其他兩群 (0-1.36%)；最後第 III 群之野生稻 (CD) 之細胞生長和形態發生功能 (2.50%) 高於其他兩群 (0.05-0.27%)，以及傳輸促進作用功能 (10.00%) 高於其他兩群 (4.89-5.29%)。由此也佐證了從對應分析雙標圖中所獲之結果。

表 5. 受稻熱病菌誘導之水稻 EST 各群在 MIPS 功能分類系統所得的功能類別統計

Table 5. Statistics of functional classification for rice EST induced by *Magnaporthe grisea* under each individual group, based on MIPS functional category system

| Functional category | | Representation (%) | | |
|---------------------|--|--|-------------------------------|-------------------|
| | | Group I (BI-1, BM-1, BQ, BI-2, CBir36, CBbare, BM-2, BU) | Group II (DN, BF-2, CX) | Group III (CD) |
| MIPS_ | Definition | | | |
| funcat ID. | | | | |
| 1 | Metabolism | 10.4 | 7.29 | 7.29 |
| 2 | Energy | 6.63 | 10.14 | 2.95 |
| 3 | Cell cycle and DNA processing | 1.16 | 2.14 | 2.36 |
| 4 | Transcription | 0.54 | 0.00 | 0.00 |
| 5 | Protein synthesis | 1.96 | 2.45 | 0.00 |
| 6 | Protein fate / folding, modification, destination | 6.41 | 6.41 | 0.45 |
| 8 | Cellular transport and transport mechanisms | 1.68 | 2.38 | 0.00 |
| 10 | Cellular communication / signal transduction mechanism | 1.48 | 0.30 | 0.00 |
| 11 | Cell rescue, defense and virulence | 1.33 | 0.88 | 0.00 |
| 13 | Ionic homeostasis | 0.40 | 0.03 | 0.00 |
| 14 | Cell growth / morphogenesis | 0.27 | 0.05 | 2.50 |
| 25 | Development (Systemic) | 0.42 | 0.37 | 0.00 |
| 30 | Control of cellular organization | 1.36 | 2.11 | 0.00 |
| 40 | Subcellular localization | 17.11 | 21.81 | 27.53 |
| 63 | Protein binding | 6.25 | 5.58 | 1.45 |
| 65 | Storage protein | 0.02 | 0.00 | 0.00 |
| 67 | Transport facilitation | 4.89 | 5.29 | 10.00 |
| 99 | Unclassified proteins | 36.83 | 35.15 | 45.45 |

討 論

抗病植物遭受病原菌侵染後之所以產生抗性，是由於活化了植物體內的抗病基因，進而透過訊號傳導鏈啟動了植物體內的防禦體制。由上述結果證實，水稻受稻熱病菌侵染後的防禦反應之過程是涉及眾多基因共同表現的一個複雜的網絡體系。但是受稻熱病菌誘導的水稻 EST 序列若是取自於混合了不同品種及不同植株部位之組織，與使用單一品種及固定植株部位組織所得的 EST 序列在基因表現上存在有明顯差異（圖 4-A, B、圖 5 及表 5），這是因為基本上不同部位組織或不同品種個體，它們的基因表現本身就是有差異的。因此，建議未來各研究單位在以 EST 為分析材料進行功能基因表現之研究時，應盡量分開品種和植株部位。

本研究結果發現，各水稻品種受稻熱病菌侵染後的次細胞定位功能之基因表現量皆會增加（表 4、圖 1），這是因為細胞中蛋白質合成後被轉運到特定的胞器中，只有轉運到正確的部位才能參與細胞的各種生命活動（如細胞週期的調控過程、細胞的訊號傳導和轉錄調控等），若細胞內的定位發生偏差，將會對細胞功能甚至生命產生重大影響（Dönnes & Höglund 2004; Zhang *et al.* 2007）。本研究進而發現受稻熱病菌誘導之野生稻 EST，以表現次細胞定位和代謝功能之基因量明

顯高於梗稻及秈稻（圖 1），此驗證了 Cho *et al.* (2004) 之結果，他指出相較於栽培稻，野生稻於營養生長期之葉片組織具有更活躍的新陳代謝能力，其表現以次細胞定位與代謝功能之基因為主。另外，Shim *et al.* (2004) 也發現在感染稻熱病之情況下野生稻的代謝功能之基因表現量超過栽培種，且指出新陳代謝變化量對於抵抗病原菌侵害扮演了不可或缺之角色。這是因為水稻在逆境狀態下為了維持細胞的動態平衡，有一些參與 ATP 合成途徑之關鍵基因會同時被誘導表現出來而活化了能量的合成代謝 (Umeda *et al.* 1994)。Song & Shi (2006) 之研究結果指出，過氧化物酶 (peroxidase, POD) 是細胞內重要的內源性活性氧清除劑，當植物受稻熱病菌入侵後，會誘發啟動某些基因的表達，合成 POD 而啟動多種代謝途徑；在植物品種與病原物互作產生親和性和非親和性反應並引發多種生理生化反應中，抗病品種的 POD 活性較感病品種上升速度快，故抗病品種啟動代謝途徑的能力相對地也較感病品種為強。而由本研究結果發現受稻熱病菌感染後，野生稻的代謝功能之基因表現量較高於梗稻及秈稻，藉此可推測野生稻啟動代謝途徑的能力較栽培稻為強，而使野生稻抵抗稻熱病菌的能力可能較高於栽培種。野生稻受稻熱病菌感染後表現出的傳輸促進作用之基因量也較高（圖 1），按可能是因為在植物抗病反應過程中，抗病訊號必須由內源訊號分子從受病菌侵入部位傳輸到整株植物，以引起相應的系統性抗性，因而較具抗性之品種傳輸內源訊號分子的速度通常也較快 (Halterman & Martin 1997)。而與細胞生長和形態發生相關的基因在野生稻表現含量也較高（圖 1），表示野生稻在對抗稻熱病菌的防禦反應過程中，細胞生長、發育過程受到較大的刺激活化，此將促使有機體能更強壯以增加防禦的能力。Cho *et al.* (2004) 指出野生稻和感染稻熱病之栽培稻所表現出的能量功能之基因量，皆較高於正常未染病的栽培種，而由本研究結果進而發現（圖 1），水稻受稻熱病菌感染後，與能量相關的基因表現量以秈稻最高，梗稻其次，而野生稻較低。另外，野生稻受稻熱病菌感染後的蛋白質循環轉變功能之基因表現量，也明顯低於栽培種（圖 1）。由此看來，未來值得針對野生稻的防禦體系做更深入探討，當有助於水稻對稻熱病抗性的分子機制研究。

受稻熱病菌誘導之水稻 EST 在不同植株部位及不同葉齡的基因表現差異，主要反映在能量和代謝功能（圖 2）：在不同生長階段中，水稻葉片於成熟期受稻熱病菌感染後的能量和代謝之功能的基因表現量最高，然後是分蘖期，幼苗期則最低甚至完全沒有表現出能量功能之基因量，按此推測水稻葉片隨著植株成長，愈到後期，其啟動代謝途徑的能力愈強，表示其對稻熱病愈具抗性。另外，本研究也發現幼苗葉片受稻熱病菌感染後的蛋白質循環轉變和次細胞定位之相關功能的基因表現量較生長後期的葉片為高，但由於幼苗期葉片的 EST 資料僅來自於 AF 基因庫，該基因庫僅有 46 條 EST 序列（表 1），此分析結果之可信度，仍待未來有更多 EST 資料加入後始能確認。目前已登錄在 dbEST 內受稻熱病菌誘導之水稻 EST，絕大多數取自葉片組織，而源自於其他不同組織者只有 BF-2 和 DN 兩個 cDNA 基因庫，但該兩者都是將不同組織混合一起（表 1），因此從現有 EST 資料的分析結果來看，最多只能確認混合組織和固定組織（葉片）間有明顯不同，尚難以真正區分出水稻在不同感染植株部位間之基因表現型式的差異性。

誌 謝

本研究承蒙行政院國家科學委員會補助經費 (NSC92-2313-B-055-006, NSC93-2313-B-055-002)，特此誌謝。

引用文獻 (Literature cited)

- Adams, M. D., M. Dubnick, A. R. Kerlavage, R. Moreno, J. M. Kelley, T. R. Utterback, J. W. Nagle, C. Fields, and J. C. Venter. 1992. Sequence identification of 2,375 human brain genes. *Nature* 355:632-634.
- Adams, M. D., J. M. Kelley, J. D. Gocayne, M. Dubnick, M. H. Polymeropoulos, H. Xiao, C. R. Merrill, A. Wu, B. Olde, R. F. Moreno, A. R. Kerlavage, W. R. McCombie, and J. C. Venter. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252:1651-1656.
- Adams, M. D., A. R. Kerlavage, R. D. Fleischmann, R. A. Fuldner, C. J. Bult, N. H. Lee, E. F. Kirkness, K. G. Weinstock, J. D. Gocayne, O. White, G. Sutton, J. A. Blake, R. C. Brandon, M. W. Chiu, R. A. Clayton, R. T. Cline, M. D. Cotton, J. Earle-Hughes, L. D. Fine, L. M. FitzGerald, W. M. FitzHugh, J. L. Fritchman, N. S. M. Geoghagen, A. Glodek, C. L. Gnehm, M. C. Hanna, E. Hedblom, P. S. Hinkle Jr., J. M. Jelley, K. M. Klimek, J. C. Kelley, L. I. Liu, S. M. Marmaros, J. M. Merrick, R. F. Moreno-Palanques, L. A. McDonald, D. T. Nguyen, S. M. Pellegrino, C. A. Phillips, S. E. Ryder, J. L. Scott, D. M. Saudek, R. Shirley, K. V. Small, T. A. Spriggs, T. R. Utterback, J. F. Weidman, Y. Li, R. Barthlow, D. P. Bednarik, L. Cao, M. A. Cepeda, T. A. Coleman, E. J. Collins, D. Dimke, P. Feng, A. Ferrie, C. Fischer, G. A. Hastings, W. W. He, J. S. Hu, K. A. Huddleston, J. M. Greene, J. Gruber, P. Hudson, A. Kim, D. L. Kozak, C. Kunsch, H. Ji, P. S. Meissner, H. Olsen, L. Raymond, Y. F. Wei, J. Wing, C. Xu, G. L. Yu, S. M. Ruben, P. J. Dillon, M. R. Fannon, C. A. Rosen, W. A. Haseltine, C. Fields, C. M. Fraser, and J. C. Venter. 1995. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* 377(6457 Suppl.):3-174.
- Andreas. R., A. Zollner, D. Maier, K. Albermann, J. Hani, M. Mokrejs, I. Tetko, U. Güldener, G. Mannhaupt, M. Münsterkötter, and H. W. Mewes. 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* 32:5539-5545.
- Baker, B., P. Zambryski, B. Staskawicz, and S. P. Dinesh-Kumar. 1997. Signaling in plant-microbe interactions. *Science* 276:726-733.
- Beh. E. J. 2004. A Bibliography of the Theory and Application of Correspondence Analysis. Vol. II-By Publication. School of Quantitative Methods and Mathematical Sciences, Univ. Western Sydney, Australia. 101 pp.
- Boguski, M. S., T. M. Lowe, and C. M. Tolstoshev. 1993. dbEST--database for "expressed sequence tags". *Nat. Genet.* 4:332-333.
- Bonaldo, M. F., G. Lennon, and M. B. Soares. 1996. Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res.* 6:791-806.
- Chiu, Y. C., M. L. Wei, C. T. Lu, and H. Y. Lu. 2007. A new procedure of functional annotation and analysis in expressed sequence tag (EST). *Crop, Enviro. Bioinfor.* 4:49-64. (in Chinese with English abstract)

- Cho, S. K., S. H. Ok, J. U. Jeung, K. S. Shim, K.W. Jung, M. K. You, K. H. Kang, Y. S. Chung, H. C. Choi, H. P. Moon, and J. S. Shin. 2004. Comparative analysis of 5211 leaf ESTs of wild rice (*Oryza minuta*). *Plant Cell Rep.* 22:839-847.
- Dönnes P., and A. Höglund. 2004. Predicting protein subcellular localization: Past, present, and future. *Genom. Proteom. Bioinform.* 2:209-215.
- Ewing, R. M., A. B. Kahla, O. Poirot, F. Lopez, and S. Audic, and J. M. Claverie. 1999. Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res.* 9:950-959.
- Feng, Q., Y. Zhang, P. Hao, S. Wang, G. Fu, Y. Huang, Y. Li, J. Zhu, Y. Liu, X. Hu, P. Jia, Y. Zhang, Q. Zhao, K. Ying, S. Yu, Y. Tang, Q. Weng, L. Zhang, Y. Lu, J. Mu, Y. Lu, L. S. Zhang, Z. Yu, D. Fan, X. Liu, T. Lu, C. Li, Y. Wu, T. Sun, H. Lei, T. Li, H. Hu, J. Guan, M. Wu, R. Zhang, B. Zhou, Z. Chen, L. Chen, Z. Jin, R. Wang, H. Yin, Z. Cai, S. Ren, G. Lv, W. Gu, G. Zhu, Y. Tu, J. Jia, Y. Zhang, J. Chen, H. Kang, X. Chen, C. Shao, Y. Sun, Q. Hu, X. Zhang, W. Zhang, L. Wang, C. Ding, H. Sheng, J. Gu, S. Chen, L. Ni, F. Zhu, W. Chen, L. Lan, Y. Lai, Z. Cheng, M. Gu, J. Jiang, J. Li, G. Hong, Y. Xue, and B. Hanand. 2002. Sequence and analysis of rice chromosome 4. *Nature* 420:316-320.
- Gene Ontology Consortium. 2006. The Gene Ontology (GO) project in 2006. *Nucleic Acids Res.* 34:D322-D326.
- Gupta, S. K., and T. C. Ghosh. 2001. Gene expressivity is the main factor in dictating the codon usage variation among the genes in *Pseudomonas aeruginosa*. *Gene* 273:63-70.
- Halterman, D. A., and G. B. Martin. 1997. Signal recognition and transduction involved in plant disease resistance. *Essays Biochem.* 32:87-99.
- Hill, M. O. 1973. Reciprocal averaging: an eigenvector method of ordination. *J. Ecol.* 61:237-249.
- Huang, X., and A. Madan. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9:868-877.
- International Rice Genome Sequencing Project. 2005. The map-based sequence of the rice. *Nature* 436:793-800.
- Jantasuriyarat, C., M. Gowda, K. Haller, J. Hatfield, G. Lu, E. Stahlberg, B. Zhou, H. Li, H. Kim, Y. Yu, R. A. Dean, R. A. Wing, C. Soderlund, and G. L. Wang. 2005. Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction. *Plant Physiol.* 138:105-115.
- Jongeneel, C. V. 2000. Searching the expressed sequence tag (EST) database: planning for genes. *Briefings Bioinform.* 1:76-92.
- Kantety, R. V. , M. L. Rota, D. E. Matthews, and M. E. Sorrells. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* 48:501-510.
- Kohchi, T., K. Fujishige, and K. Ohyama. 1995. Construction of an equalized cDNA library from *Arabidopsis thaliana*. *Plant J.* 8:771-776.

- Lepš, J., and P. Šmilauer. 1999. *Multivariate Analysis of Ecological Data*. Faculty of Biological Science. Univ. South Bohemia, Ceske Budejovice. 110 pp.
- Liang, M. 2005. Review of researches on inheritance and breeding of blast resistance in rice. *Chinese Agric. Bull.* 21:341-345. (in Chinese with English abstract)
- Liu, M. Y., Y. J. T. Zeng, and C. N. Sun. 1981. Diamondback moth resistant to several synthetic pyrethroids. *J. Ecol. Entomol.* 74:393-396.
- Lu, H. Y., C. T. Chen, Y. C. Chiu, C. T. Lu, and M. L. Wei. 2006. An automated system for EST functional analysis and its application in rice genes. p.77-78. *in the Bioinformatics in Taiwan Symposium and Workshop*. Bioinformatics Soc. Taiwan. Taichung.
- Mewes, H. W., D. Frishman, K. F. X. Mayer, M. Münsterkötter, O. Noubibou, P. Pagel, T. Rattei, M. Oesterheld, A. Ruepp, and V. Stümpflenand. 2006. MIPS: analysis and annotation of proteins from whole genomes in 2005. *Nucleic Acids Res.* 34:169-172.
- Okasanen, J. 2004. *Multivariate Analysis in Ecology-Lecture Notes*. p.63-70. Department of Biology, Univ. Oulu, Oulu, Finnish.
- Polymeropoulos, M. H., H. Xiao, J. M. Sikela, M. Adams, J. C. Venter, and C. R. Merril. 1993. Chromosomal distribution of 320 genes from a brain cDNA library. *Nat. Genet.* 4:381-386.
- Rauyaree, P., W. Chol, E. Fang, B. Blackmon, and R. A. DEAN. 2001. Gene expressed during early stages of rice infection with the rice blast fungus *Magnaporthe grisea*. *Mol. Plant Pathol.* 2:347-354.
- Rounsley, S. D., A. Glodek, G. Sutton, M. D. Adams, C. R. Somerville, J. C. Venter, and A. R. Kerlavage. 1996. The construction of *Arabidopsis* expressed sequence tag assemblies. A new resource to facilitate gene identification. *Plant Physiol.* 112:1177-1183.
- Rudd, S. 2003. Expressed sequence tags: alternative or complement to whole genome sequence? *Trends Plant Sci.* 8:321-329.
- Sasaki, T., T. Matsumoto, K. Yamamoto, K. Sakata, T. Baba, Y. Katayose, J. Wu, Y. Niimura, Z. Cheng, Y. Nagamura, B. A. Antonio, H. Kanamori, S. Hosokawa, M. Masukawa, K. Arikawa, Y. Chiden, M. Hayashi, M. Okamoto, T. Ando, H. Aoki, K. Arita, M. Hamada, C. Harada, S. Hijishita, M. Honda, Y. Ichikawa, A. Idonuma, M. Iijima, M. Ikeda, M. Ikeno, S. Ito, T. Ito, Y. I. Ito, A. Iwabuchi, K. Kamiya, W. Karasawa, S. Katagiri, A. Kikuta, N. Kobayashi, I. Kono, K. Machita, T. Maehara, H. Mizuno, T. Mizubayashi, Y. Mukai, H. Nagasaki, M. Nakashima, Y. Nakama, Y. Nakamichi, M. Nakamura, N. Namiki, M. Negishi, I. Ohta, N. Ono, S. Saji, K. Sakai, M. Shibata, T. Shimokawa, A. Shomura, J. Song, Y. Takazaki, K. Terasawa, K. Tsuji, K. Waki, H. Yamagata, H. Yamane, S. Yoshiki, R. Yoshihara, K. Yukawa, H. Zhong, H. Iwama, T. Endo, H. Ito, J. H. Hahn, H. I. Kim, M. Y. Eun, M. Yano, J. Jiang, and T. Gojoboriand. 2002. The genome sequence and structure of rice chromosome 1. *Nature* 420:312-316.
- Shim, K. S., S. K. Cho, J. U. Jeung, K. W. Jung, M. K. You, S. H. Ok, Y.S. Chung, K. H. Kang, H. G. Hwang, H. C. Choi, H. P. Moon, and J. S. Shin. 2004. Identification of fungal (*Magnaporthe grisa*) stress-induced genes in wild rice (*Oryza minuta*). *Plant Cell Rep.* 22:599-607.

- Song, H. C., and X. Q. Shi. 2006. The relationships of changes in antioxidant enzymes of rice seedling infected by *Magnaporthe grisea* to the blast resistance. Nat. Sci. J. Hainan Univ. 24:378-382. (in Chinese with English abstract)
- Statsoft Inc. 2002. STATISTICA: The Small Book Chinese Version. USA. 144 pp.
- Tan, Q., K. Brusgaard, T. A. Kruse, E. Oakeley, B. Hemmings, H. Beck-Nielsen, L. Hansen, and M. Gaster. 2004. Correspondence analysis of microarray time-course data in case-control design. J. Biomed. Inform. 37:358-365.
- Tekaia, F., Y. Edouard, and D. Bernard. 2002. Amino acid composition of genomes, lifestyles of organisms, and evolutionary trends: a global picture with correspondence analysis. Gene 297:51-60.
- Tyagi, A. K., J. P. Khurana, S. Khurana, A. Gaur, A. Kapur, V. Gupta, D. Kumar, V. Ravi, S. Vij, P. Khurana, and S. Sharma. 2004. Structural and functional analysis of rice genome. J. Genet. 83:79-99.
- Umeda, M., C. Hara, Y. Matsubayashi, H. H. Li, Q. Liu, F. Tadokoro, S. Aotsuka, and H. Uchimiya. 1994. Expressed sequence tags from cultured cells of rice (*Oryza sativa* L.) under stressed conditions: analysis of transcripts of genes engaged in ATP-generating pathways. Plant Mol. Biol. 25:469-478.
- Wei, M. L., C. T. Lu, and H. Y. Lu. 2005a. Research and development of expressed sequence tag (EST) in cereal crops. Sci. Agric. 53(1,2):15-21. (in Chinese)
- Wei, M. L., H. Y. Lu, and C. T. Lu. 2005b. Bioinformatics collection and statistics of rice disease / pest resistance genomics research. Crop Environ, Bioinform. 2:295-306. (in Chinese with English abstract)
- Zhang, S., B. Huang, X. F. Xia, and Z. R. Sun. 2007. Bioinformatics research in subcellular localization of proteins. Progress Biochem. Biophys. 34:573-579. (in Chinese with English abstract)

Comparative Analysis of Expressed Sequence Tag (EST) in Rice Induced by *Magnaporthe grisea*¹

Hsiu-Ying Lu^{2,4}, Yi-Chia Chiu³, Cheng-Tao Chen²,
Meng-Li Wei² and Chun-Tang Lu²

Abstract

Lu, H. Y., Y. C. Chiu, C. T. Chen, M. L. Wei, and C. T. Lu. 2007. Comparative analysis of expressed sequence tag (EST) in rice induced by *Magnaporthe grisea*. J. Taiwan Agric. Res. 56:261-280.

Rice blast is caused by fungus pathogen *Magnaporthe grisea*, which is a major destructive disease in rice production worldwide. Expressed sequence tag (EST) is the fragment of cDNA sequence. Direct study of EST is therefore helpful to obtain the gene expression information. In order to investigate the differences of gene expression patterns of rice among varieties, tissues and developmental stages after rice-blast infection, we made a comparative analysis of rice ESTs induced by *M. grisea* from 14 cDNA libraries in NCBI-dbEST database. For each individual library, EST functional annotation, classification and representation statistics were performed. Clustering analysis and correspondence analysis were also used to detect the differences in gene expression patterns among libraries and identify the functional categories involved. Results showed that wild rice infected by *M. grisea* exhibited more transcripts involved in metabolism, cell growth / morphogenesis, subcellular localization and transport facilitation compared to indica rice and japonica rice. The representation of genes associated with metabolism and energy functions in infected rice leaf at mature stage was highest, followed by tillering stage, seedling stage. When rice ESTs induced by *M. grisea* were mixtures of sequences with different rice varieties and tissues, it appeared diversity of gene expression from those of rice ESTs from single source. These findings will facilitate understanding of the resistance mechanisms against rice blast.

Key words: Rice, Rice blast, Expressed sequence tag, Functional annotation and classification, Statistics of gene representation, Clustering analysis, Correspondence analysis.

-
1. Contribution No.2301 from Agricultural Research Institute, Council of Agriculture. Accepted: November 29, 2007.
 2. Respectively, Senior Researcher, Project Assistant, Assistant Researcher and Assistant Researcher, Crop Science Division, ARI, Wufeng, Taichung, Taiwan, ROC.
 3. Master, Graduate Institute of Biotechnology, Chaoyang University of Technology, Wufeng, Taichung, Taiwan ROC.
 4. Corresponding author, e-mail:iying@wufeng.tari.gov.tw; Fax:(04)23390528.