

II. 6 台灣作物種源庫記錄之電腦化作業

劉 清

台灣省農業試驗所

前 言

種源庫 (Germplasm bank) 資料對遺傳及育種工作來說十分重要，從事品種改良工作者，無法憑空產生新品種，而經常是由種源庫中之各種不同品種，按其育種目的，選擇最適合之材料來進行品種改良工作。換言之，科學家直接利用種源庫的資料，而間接利用種源庫的遺傳材料。一個收集材料再多再豐富的種源庫，如果沒有一套完整的資料處理系統，對科學家來說毫無用處。因此在種源庫設立之初，即應對材料的收集 (Collection)，形態及農藝特性 (Morpho-agronomic characters) 的調查及遺傳特性的評估 (Evaluation) 作業等建立一完整的資料記錄，處理及取出系統 (Data recording, processing and retrieval system)。此類系統的設計不只要針對個別種源庫之需要，設備，尚要考慮到能與其它有關種源庫之資料系統相互連接溝通。所記錄之資料基本上要包括一

般資料 (General information)，如品種名稱、來源等等；形態及農藝特性之調查記錄，遺傳特性之評估結果，材料保存數量及與其它機構之交換記錄等資料。

電腦化處理對種源庫資料系統的作業來說，並非絕對必要，但對大型種源庫而言卻是不可或缺。經由電腦的利用，科學家可極有效率的儲存，取出及利用數量龐大的資料。然而電腦本身無法完成任何工作，需由有經驗之電腦程式設計及管理人員 (Computer programming and management personels) 來設計資料系統之程式作業。因此種源庫資料之電腦化作業是由科學家及電腦專家共同合作完成，科學家決定那些資料必須加以記錄保存，電腦專家則可幫助科學家決定最有效率的資料記錄，處理及取出方法。

資料的種類

種源庫由各地收集到的材料首先需在某些特定地點種植繁殖，一方面增加保有之材料數量，另方面可有系統的記錄其形

態及農藝特性，並作初步之評估，以進一步篩選有利用及保存價值之材料。這些特性記錄及評估結果最後均應導致有效的資料記錄及處理方式，並能迅速的重行取出科學家所需要的資料。

基本上，種源庫所記錄之資料可分為以下數類⁽²⁾：

1. General information

Variety or accession number

Collection information (collector, date, etc.)

Nomenclature information (scientific and common)

Origin (country, state or province, and precise locality)

Storage information (locations, date placed in storage, etc.)

Distribution information

2. Environmental information

Altitude

Latitude

Longitude

Climatic

Ecological

Soil description

3. Organismic information

Morpho-agronomic

Genetic

Physiological

Biochemical

Environmental damage (cold, drought, wind or other susceptibility)

Pest and disease information (bacterial, fungal, viral, insect, nematode)

Rejuvenation (viability of seeds, time in storage)

4. Use information

Including breeding, genetics, performance tests, etc.

5. Food characteristics

6. Industrial characteristics

7. Bibliographic information

各種作物種源庫之資料大多可用上法整理。當然並非所有種源庫之資料系統都必需具有以上各類資料，各作物種源庫可視其需要增減資料種類。若條件許可，則資料種類愈多愈完整愈好，事實上，一個理想之種源庫資料系統，應能對遺傳、育種、生理、病蟲害及生化等各方面之研究提供有價值的參考資料。

由於不同作物之資料記錄亦不同，因此各有關專家應對各作物分別設計理想方便的特性記錄表格，表格之設計愈理想、方便，則記錄亦愈精確，同時亦可減少在

coding 時容易發生之錯誤。

資料的表示

資料一般以“文數字”(Alphanumeric characters)存放於記錄媒體上(如打孔卡片、磁帶或磁碟)。所謂文數字是由26個英文字母(A到Z),10個阿拉伯數字(0到9)及一些其它特殊符號(如+、-、*、/等)組成。以文數字表示之代號系統稱為Alphanumeric Coding System,簡稱ACS。以文數

字來存放資料的好處是定性或定量的資料皆可容易的表示(對人類而言),如綠色以英文字母G或G_n表示,一個性狀之出現與否可用+,一符號代表,株高或穗數即以實際數據表示,但採用ACS的缺點之一是以電腦來處理資料時的速度較慢,尤其在作Sorting或Comparison時為然。另外以文數字記錄資料有時不易表示二個記錄間之關係,如抗病性可用以下二種表示法:

Code	Reaction	Code	Reaction
1 =	Immune	I =	Immune
3 =	Resistant	R =	Resistant
5 =	Tolerant	T =	Tolerant
7 =	Hypersensitive	H =	Hypersensitive
9 =	Susceptible	S =	Susceptible

左邊之數字代號可表示出各種反應之相對程度,但右邊之文字代號則不能。

由於ACS有以上各種資料處理上的缺點,因此目前一些大種源庫(如IRRI)已改採用Decimal Coding System,簡稱DCS⁽¹⁾。除了名字以外之所有記錄皆以數字表示。採用DCS的優點及缺點正好與ACS之優點及缺點相反。DCS之資料在電腦處理時非常迅速,且由於採用數值制,因此記錄間之關係很容易了解,如以上所舉對病害反應之數字表示法。DCS對有關量的資料,其記錄

法與ACS一樣,即以實際數據表示,但對如顏色、抗病性、抗蟲性、耐旱性、耐寒性等資料則以0到9等10個阿拉伯數字來表示由少到多,由淺到深或由最好到最差之相對等級,另外以空白(Blank)來代表沒有資料。採用DCS還有另外一個好處即是資料在記錄媒體上的存放非常經濟,只要對各項資料設計出一種數字代號系統,則所有資料皆可用簡單數字表示。目前世界各國對各類資料的代號系統尚未統一,但新成立種源庫之資料代號系統應與其它關係接近之大種源庫代號系統一致

或接近。如此可有利於各種源庫之資料流通。一種代號系統之設計應仔細謹慎，考慮周詳，且不宜時常更改。

此外，資料在記錄媒體上的存放規格及代號系統之設計，最好也能與目前一般通用之各種 Computer packages（如 Statistical Analysis System，簡稱 SAS，或 Statistical Package for the Social Science，簡稱 SPSS，等）相配合。此類 Packages 在應用上非常方便，包括有一般常用之資料整理及統計分析程式，且各大電腦中心皆裝置有一或數套此類 Packages。種源庫資料之記錄方法如能與這些 Packages 之要求相配合，則資料無需經過特殊轉換即可直接利用這些 Packages 作許多資料整理及分析工作，可節省許多花在程式設計上的時間及人力。

農業試驗所稻米種源之資料處理系統

農試所稻米種源庫資料之電腦化作業目前已初步設計完成。雖由於受到農試所小型電腦設備（HP 9825 系統）的限制，無法充分發揮一個電腦化資料處理系統的功能，但經由精密的程式設計，已能達到種源庫資料電腦化處理的首要目標，即資料“存”、“取”之自動化。此外農試所稻米種源庫電腦資料之記錄規格與農試所

統計程式系統所要求之輸入規格均一致，因此亦可直接用來作其它統計分析，如頻度分布，迴歸及相關分析等等。

目前農試所仍採用 ACS，計劃在未來改採 DCS。農試所保存有將近三千種稻米品種，各品種分別調查 19 種農藝性狀及 7 種遺傳評估性狀，26 種特性記錄連同品種代號皆記錄於磁帶及磁碟上，並分別建立一個主檔（Master file）及一個異動檔（Transaction file），記錄媒體上之資料可隨時修改更新。根據以上之資料檔，農試所已發展出一套“種源庫資料取出程式”（Germplasm bank information retrieval program，簡稱 GBIRP）其目的在找出具有某些特性之品種代號及記錄。研究者使用此一程式只需在一張請求表格（Request form）上填上其所指定之特性及每一特性之指定變異範圍，則 GBIRP 即可根據請求表格上的要求，在所有記錄品種中尋出符合研究者要求之品種。如圖 1 之請求表格上，研究者所要求的品種為(1)台灣梗稻（農試所種源庫，品種代號 1 至 713 為台灣梗稻），(2)二期作株高在 100 到 110 公分間，(3)穗重在 3 g 以上，(4)葉稻熱病之罹病程度在中感（等級為 3）以下，(5)直立不倒伏，(6)穀粒長寬比在 3.0 以下。圖 2 即為依照圖 1 請求表格上之條件所選出之合格品種，所需時間約為 2 分 11 秒。

圖 1：研究者使用“種源庫資料取出程式”之請求表格例樣。

TAIWAN AGRICULTURAL RESEARCH INSTITUTE
INFORMATION RETRIEVAL REQUEST FORM
(RICE)

RESEARCHER _____ DEPARTMENT _____
DATE SUBMITTED _____ DATE NEEDED _____

RECORD	CRITERION(1)		CRITERION(2)	
	OPERATOR	VALUE	OPERATOR	VALUE
1. VARIETY NUMBER				
2. PLANT HEIGHT OF 1ST CROP(cm)				
3. PLANT HEIGHT OF 2ND CROP(cm)	>	100	<	110
4. NUMBER OF TILLERS OF 1ST CROP				
5. NUMBER OF TILLERS OF 2ND CROP				
6. COLOR OF LEAF BLADE				
7. COLOR OF BASAL PORTION OF CULM				
8. WEIGHT OF PANICLE(g)	>	3.0		
9. AWN				
10. COLOR OF STIGMA				
11. COLOR OF APICULUS AT BLOOMING STAGE				
12. COLOR OF APICULUS AT MATURITY				
13. COLOR OF LEMMA				
14. SEEDLING BLAST				
15. LEAF BLAST	<	4		
16. NECK BLAST(PERCENTAGE)				
17. SHEATH BLIGHT SPOT(PERCENTAGE)				
18. SHEATH BLIGHT SPOT(DEGREE)				
19. LODGING RESISTANCE	=	1		
20. DROUGHT RESISTANCE				
21. RESPONSE TO NITROGEN				
22. PERCENTAGE OF SHATTERING				
23. DATE OF HEADING OF 1ST CROP				
24. DATE OF HEADING OF INTERMEDIATE CROP				
25. DATE OF HEADING OF 2ND CROP				
26. RATIO OF LENGTH TO WIDTH OF GRAIN	≤	3.0		
27. ENDOSPERM				

OPERATOR: =EQUAL, > GREATER THAN, < LESS THAN, ≠ NOT EQUAL TO,
>= GREATER THAN OR EQUAL TO, ≤ LESS THAN OR EQUAL TO.

圖 2 : 應圖 1 之要求 “種源庫資料取出程式” 所印出之報表 :

REC NO

RECORD

-
- 1 = VARIETY NUMBER
 - 2 = PLANT HEIGHT OF 1ST CROP(cm)
 - 3 = PLANT HEIGHT OF 2ND CROP(cm)
 - 4 = NUMBER OF TILLERS OF 1ST CROP
 - 5 = NUMBER OF TILLERS OF 2ND CROP
 - 6 = COLOR OF LEAF BLADE
 - 7 = COLOR OF BASAL PORTION OF CULM
 - 8 = WEIGHT OF PANICLE(g)
 - 9 = AWN
 - 10 = COLOR OF STIGMA
 - 11 = COLOR OF APICULUS AT BLOOMING STAGE
 - 12 = COLOR OF APICULUS AT MATURITY
 - 13 = COLOR OF LEMMA
 - 14 = SEEDLING BLAST
 - 15 = LEAF BLAST
 - 16 = NECK BLAST (PERCENTAGE)
 - 17 = SHEATH BLIGHT SPOT (PERCENTAGE)
 - 18 = SHEATH BLIGHT SPOT (DEGREE)
 - 19 = LODGING RESISTANCE
 - 20 = DROUGHT RESISTANCE
 - 21 = RESPONSE TO NITROGEN
 - 22 = PERCENTAGE OF SHATTERING
 - 23 = DATE OF HEADING OF 1ST CROP
 - 24 = DATE OF HEADING OF INTERMEDIATE CROP
 - 25 = DATE OF HEADING OF 2ND CROP
 - 26 = RATIO OF LENGTH TO WIDTH OF GRAIN
 - 27 = ENDOSPERM

PLEASE REFER TO 'A MONOGRAPH OF RICE VARIETIES
PRESERVED BY TAIWAN AGRICULTURAL RESEARCH INSTITUTE
FOR SYMBOLS USED FOR VARIOUS CHARACTERS

THESE ARE VARIETIES WITH:

RECORD(3)	>	100
RECORD(3)	<	110
RECORD(8)	>	3.0
RECORD(15)	<	4
RECORD(19)	=	
RECORD(26)	<=	3.0

圖 2 續

REC.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
(66)	90	107.6	100.2	11.2	13.4	DGn	Gn	3.34	—	W	LGn	O	YS	S3
(68)	92	115.4	106.2	12.4	10.0	DGn	Gn	3.16	—	W	LGn	O	YS	S1
(74)	98	112.6	101.7	11.0	10.6	DGn	Gn	3.74	—	W	LGn	X	YS	S1
(76)	100	111.0	101.6	12.0	10.4	DGn	Gn	3.64	—	W	LGn	O	YS	S2
(81)	105	103.6	105.0	9.4	10.0	DGn	Gn	3.90	—	W	LGn	O	YS	S3
(94)	129	109.6	107.6	12.4	12.4	DGn	Gn	3.50	—	W	LGn	X	YS	S3
(152)	193	113.0	103.2	13.2	10.6	DGn	Gn	3.18	—	W	LGn	X	YS	R3
(201)	262	111.0	104.7	10.2	10.0	DGn	Gn	4.58	—	W	LGn	O	YS	S3
(240)	353	111.4	104.0	11.2	10.6	DGn	Gn	3.70	RS	W	LGn	O	YS	S2
(248)	361	116.8	109.0	13.2	12.4	DGn	Gn	4.30	RM	W	LGn	X	YS	S3
(293)	448	116.6	103.4	14.8	12.6	DGn	Gn	3.54	—	W	LGn	X	YS	S3

15	16	17	18	19	20	21	22	23	24	25	26	27
3	100	56.8	22.5		—	D	12.4	6.10		10.11	2.09	G 1
3	5	57.8	11.1		++	C	21.5	6.16		10.13	2.11	G 1
3	10	63.3	26.6		—	D	10.8	6.10		10.3	2.07	G 1
3	65	43.9	13.1		+	C	12.3	6.13		10.4	1.94	G 1
3	20	45.3	11.8		+	D	6.2	6.7		10.7	2.00	G 1
3	10	63.6	14.7		+	D	8.0	6.11		10.12	2.01	G 1
3	100	52.8	12.0		d	D	2.5	6.13		10.11	2.18	G 1
3	40	34.3	8.7		++	C	5.2	6.10		10.9	2.08	G 1
3	50	29.9	15.5		+	D	7.6	6.16		10.10	2.02	G 1
3	20	25.1	6.9		+	D	3.7	6.12		10.13	2.10	G 1
3	0	33.0	14.5		+	C	12.3	6.12		10.12	1.97	G 1

***** END OF RETRIEVAL *****

參考文獻

1. Chang, T. T. 1976. Manual on genetic conservation of rice germplasm for evaluation and utilization. IRRI, Los Banos, Philippines.
2. Rogers, D. J., B. Snoad and L. Seidewitz. 1975. Documentation for genetic resources centers. Pages 399-405 in O. H. Frankel and J. G. Hawkes, ed. Crop genetic resources for today and tomorrow. Cambridge Univ. Press, Cambridge.