Bot. Bull. Academia Sinica 28: 219-235 (1987)

STUDIES ON THE DYNAMIC MODEL OF PLANT ADAPTATION OF QUANTITATIVE CHARACTERS

I. Estimation of Space-Time Parameters[1,2]

HISU-YING LU and HONG-PANG WU

Institute of Botany, Academia Sinica
Nankang, Taipei, Taiwan 11529, Republic of China

(Received February 20, 1987; Accepted May 15, 1987)

Abstract

Linear regression model is useful to evalue the adaptation of a quantitative character to the change of environments and can be applied to many cases. However, the quantitative character of a plant often varies with time in different stages of growth and development. This results in the diversity of adaptation for different growing periods. Therefore, a dynamic model was established to integrate the time factor into the linear model to elucidate the adaptation of genotypes. Twenty-four genetical lines of ***Arabidopsis thaliana*** were used to demonstrate the empirical model and to test its reliability and applicability.

Based on the results that the response of each genotype across environments for different growing period can be represented approximately by a straight line, the dynamic model was appropriate to assess the relative degree of adaptation among genotypes. Moreover, a procedure which combines cluster method and the linear regression was applied for sufficiently grouping genotypes into homogeneous subsets within which all genotypes share a common regression and have no difference between phenotypic mean. It provides information on both similarity and adaptation. Comparisons among response of different genotypes may be helpful to selection in the plant varietal trails. Three kinds of stability indices, i.e. growth stability index, environment stability index and growth × environment stability index, for each genotype were yielded in the dynamic model of this study. The results also indicated independence between growth stability and environment stability, and significant positive correlations between the phenotypic mean and stability indices.

Key words: Dynamic model; genotype-time-environment interaction; linear regression; adaptation; cluster method; similarity; ***Arabidopsis thaliana***; phenotypic stability.

Introduction

Regression methods for studying genotype-environment (G × E) interaction

[1] Paper No. 328 of the Scientific Journal Series, Institute of Botany, Academia Sinica
[2] This work was supported by National Science Council Grant NSC No. 75-0201-B001-14.

are widely used to investigate the response patterns of genotypes. A number of papers invloving such analyses have been published. Typically, the linear regression of genotypic performance in different environments on an environmental index is useful to evalue the relative degree of adaptation among genotypes. Genotypes with a regression coefficient b near zero (Bucio Alanis *et al.*, 1966; Perkins and Jinks, 1968; Lu and Wu, 1986), or with a $(1+b)$ near 1.0 (Yates and Cochran, 1938; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966) and a high mean performance are regarded as being well adapted to all environments. On the other hand, genotypes which are unstable or plastic will have a b value different from zero. Such a relationship facilitates the decision-making process in a breeding programme. In the absence of non-allelic interactions, this approach can be equally useful in predicting performance over generations.

However, the quantitative character of plant often varies with time when it is at different stages of growth and development. This results in the diversity of adaptation for different growing periods. At present, little or none of the effort is concentrated on the question of how to assess G × E interaction in developmental processes of plant. Therefore, from the viewpoint of linear regression, a dynamic model will be established by inducting time factor into the linear model to elucidate the adaptation of genotypes.

The objective of the regression approach in the study of G × E interaction is not in the point estimation of the coefficients, but rather the presentation of empirical comparisons among responses of different genotypes. When the number of genotypes is small, a satisfactory solution may be attainable by inspecting the individual regression lines. However, this is too laborious if the number of genotypes is large. Thus, grouping must be done by cluster analysis. As we have known, when using cluster analysis to elucidate G × E interaction, many different dissimilarity measures and clustering strategies have been used. Particular choices of these can result in different cluster groups and the acceptance or rejection of any particular choice may be difficult to justify. Many research workers (e.g., plant breeders) believe that the effective partition should enjoy certain optimal properties: The number of groups should be as small as possible, while maximum homogeneity within groups and good separation between the groups are desirable. For obtaining such a partition, Lin and Thompson (1975) extended the regression approach by systematically grouping genotypes into homogeneous subsets within which all genotypes share a common regression and have no difference between phenotypic means. It provides complete information on both similarity and adaptation. The grouping based on the critical F-value as a stopping criterion can avoid heterogeneity or regression line within groups. This kind of grouping will make it easier for researchers to identify the similar response patterns of genotypes.

By means of these approaches, the purpose of the present study is to establish

the dynamic model for analysing G × E interaction and to extend it to include the grouping of genotypes. Twenty-four genetical lines of *Arabidopsis thaliana* were used to demonstrate the empirical model and to test its reliability and applicability.

Statistical Models

Suppose the performance of n genotypes in p environments is measured at m different stages of growth for q replicates. Let y_{tijk} represents the observed performance of the i^{th} ($i = 1, \ldots, n$) genotype in the j^{th} ($j = 1, \ldots, p$) environment for the k^{th} ($k = 1, \ldots, q$) replicate at the time t ($t = 1, \ldots, m$). The basic model is

$$y_{tijk} = \mu + T_t + G_i + E_j + \alpha_{ti} + \beta_{tj} + g_{ij} + \gamma_{tij} + e_{tijk} \tag{1}$$

where μ is the overall mean; T_t is the t^{th} time effect; G_i is the i^{th} genotypic effect; E_j is the j^{th} environmental effect; $\{\alpha_{ti}\}$, $\{\beta_{tj}\}$, $\{g_{ij}\}$, $\{\gamma_{tij}\}$ are G × T, T × E, G × E and G × T × E interactions, respectively; and the error term e_{tijk} is independently normal with mean of zero and variance σ^2.

The model used by Finlay and Wilkinson (1963) and by Perkins and Jinks (1968) involves the regression of g_{ij} on E_j effects. If, in this model, there are also linear relationship between α_{ti} and T_t, and between γ_{tij} and β_{tj}, respectively. That is, we can obtain

$$\begin{aligned} \alpha_{ti} &= \xi_i T_t + \eta_{ti} \\ g_{ij} &= b_i E_j + \delta_{ij} \\ \gamma_{tij} &= \phi_i \beta_{tj} + \theta_{tij} \end{aligned} \tag{2}$$

where ξ_i, b_i and ϕ_i are the regression coefficients, and η_{ti}, δ_{ij}, θ_{tij} represent the residual deviations from the three fitted regressions, respectively. Such regression analyses are expressed in a way analogous to Finlay-Wilkinson type of analysis, therefore, equation (1) can be written as

$$\begin{aligned} y_{tijk} = \mu &+ (1 + \xi_i) T_t + G_i + (1 + b_i) E_j \\ &+ (1 + \phi_i) \beta_{tj} + \eta_{ti} + \delta_{ij} + \theta_{tij} + e_{tijk} \end{aligned} \tag{3}$$

with the side conditions

$$\sum_{t=1}^{m} T_t = 0, \quad \sum_{i=1}^{n} G_i = 0, \quad \sum_{j=1}^{p} E_j = 0, \quad \sum_{t=1}^{m} \beta_{tj} = \sum_{j=1}^{p} \beta_{tj} = 0,$$

$$\sum_{i=1}^{n} \xi_i = 0, \quad \sum_{t=1}^{m} \eta_{ti} = \sum_{i=1}^{n} \eta_{ti} = 0,$$

$$\sum_{i=1}^{n} b_i = 0, \quad \sum_{i=1}^{n} \delta_{ij} = \sum_{j=1}^{p} \delta_{ij} = 0,$$

$$\sum_{i=1}^{n} \phi_i = 0, \quad \sum_{t=1}^{m} \theta_{tij} = \sum_{i=1}^{n} \theta_{tij} = \sum_{j=1}^{p} \theta_{tij} = 0.$$

Least squares estimates of the parameters under the above constraints are given by

$$\begin{cases} \hat{\mu} = \bar{y}.... \\ \hat{T}_t = \bar{y}_{t...} - \bar{y}.... \\ \hat{G}_i = \bar{y}_{.i..} - \bar{y}.... \\ \hat{E}_j = \bar{y}_{..j.} - \bar{y}.... \\ \hat{\alpha}_{ti} = \bar{y}_{ti..} - \bar{y}_{t...} - \bar{y}_{.i..} + \bar{y}.... \\ \hat{\beta}_{tj} = \bar{y}_{t.j.} - \bar{y}_{t...} - \bar{y}_{..j.} + \bar{y}.... \\ \hat{g}_{ij} = \bar{y}_{.ij.} - \bar{y}_{.i..} - \bar{y}_{..j.} + \bar{y}.... \\ \hat{\gamma}_{tij} = \bar{y}_{tij.} - \bar{y}_{ti..} - \bar{y}_{t.j.} - \bar{y}_{.ij.} + \bar{y}_{t...} + \bar{y}_{.i..} + \bar{y}_{..j.} - \bar{y}.... \\ \hat{\xi}_i = \sum_{t=1}^{m} \hat{\alpha}_{ti} \hat{T}_t \Big/ \sum_{t=1}^{m} \hat{T}^2_i \\ \hat{b}_i = \sum_{j=1}^{p} \hat{g}_{ij} \hat{E}_j \Big/ \sum_{j=1}^{p} \hat{E}^2_j \\ \hat{\phi}_i = \sum_{t=1}^{m} \sum_{j=1}^{p} \hat{\gamma}_{tij} \hat{\beta}_{tj} \Big/ \sum_{t=1}^{m} \sum_{j=1}^{p} \hat{\beta}^2_{ij} \end{cases} \quad (4)$$

The analysis of variance is shown in Table 1. Since we are interested individually in the fluctuations of adaptation of n genotypes grown in p environments during the growth periods, the main effects, T, G and E are considered to be fixed. The expected mean squares shown in the table are calculated in the usual way. In the dynamic model, we are confining our attention to that the proportion of variation of the i^{th} genotype over environments throughout the whole growth period is accounted for by fitting a regression slope. From the expected mean squares, the appropriate statistic for testing hypothesis

H_0: $\xi_1 = \xi_2 = \ldots = \xi_n = 0$ is

$$F_0 = M_\xi / M_\eta$$

which is distributed as $F_{(n-1),\,(n-1)(m-2)}$. For testing hypothesis

H_0: $b_1 = b_2 = \ldots = b_n = 0$, the test statistic is

$$F_0 = M_b / M_\delta$$

which is distributed as $F_{(n-1),\,(n-1)(p-2)}$. To test hypothesis

H_0: $\phi_1 = \phi_2 = \ldots = \phi_n = 0$, we could use the statistic

$$F_0 = M_\phi / M_\theta$$

which is distributed as $F_{(n-1),\,(n-1)(pm-p-m)}$.

If the heterogeneity between regressions mean square (M.S.) is significant when compared with residual M.S., the predictions of interactions based on the linear regression will have considerable practical value. If both the heterogeneity between regressions M.S. and residual M.S. are significant, it indicates that

Table 1. *ANOVA*

Source	D.F.	M.S.	S.S.
Time (T)	$m-1$	M_T	$npq\sum_{t=1}^{m}(\bar{y}_{t\cdots}-\bar{y}_{\cdots})^2$
Environment (E)	$p-1$	M_E	$mnq\sum_{j=1}^{p}(\bar{y}_{\cdot\cdot j\cdot}-\bar{y}_{\cdots})^2$
Genotype (G)	$n-1$	M_V	$mpq\sum_{i=1}^{n}(\bar{y}_{\cdot i\cdot\cdot}-\bar{y}_{\cdots})^2$
T×E	$(m-1)(p-1)$	M_β	$nq\sum_{t=1}^{m}\sum_{j=1}^{p}(\bar{y}_{t\cdot j\cdot}-\bar{y}_{t\cdots}-\bar{y}_{\cdot\cdot j\cdot}+\bar{y}_{\cdots})^2$
G×T	$(n-1)(m-1)$	M_α	$pq\sum_{t=1}^{m}\sum_{i=1}^{n}(\bar{y}_{ti\cdot\cdot}-\bar{y}_{t\cdots}-\bar{y}_{\cdot i\cdot\cdot}+\bar{y}_{\cdots})^2$
Het. bet. reg.'s	$n-1$	M_ξ	$pq\sum_{i=1}^{n}[\sum_{t=1}^{m}(\bar{y}_{ti\cdot\cdot}-\bar{y}_{t\cdots}-\bar{y}_{\cdot i\cdot\cdot}+\bar{y}_{\cdots})(\bar{y}_{t\cdots}-\bar{y}_{\cdots})]^2/\sum_{t=1}^{m}(\bar{y}_{t\cdots}-\bar{y}_{\cdots})^2$
Residual	$(n-1)(m-2)$	M_η	$SS_\alpha-SS_\xi$
G×E	$(n-1)(p-1)$	M_g	$mq\sum_{i=1}^{n}\sum_{j=1}^{p}(\bar{y}_{\cdot ij\cdot}-\bar{y}_{\cdot i\cdot\cdot}-\bar{y}_{\cdot\cdot j\cdot}+\bar{y}_{\cdots})^2$
Het. bet. reg.'s	$n-1$	M_b	$mq\sum_{i=1}^{n}[\sum_{j=1}^{p}(\bar{y}_{\cdot ij\cdot}-\bar{y}_{\cdot i\cdot\cdot}-\bar{y}_{\cdot\cdot j\cdot}+\bar{y}_{\cdots})^2(\bar{y}_{\cdot\cdot j\cdot}-\bar{y}_{\cdots})]^2/\sum_{j=1}^{p}(\bar{y}_{\cdot\cdot j\cdot}-\bar{y}_{\cdots})^2$
Residual	$(n-1)(p-2)$	M_δ	SS_g-SS_b
G×T×E	$(n-1)(m-1)(p-1)$	M_τ	$q\sum_{i=1}^{n}\sum_{t=1}^{m}\sum_{j=1}^{p}(\bar{y}_{tij\cdot}-\bar{y}_{ti\cdot\cdot}-\bar{y}_{t\cdot j\cdot}-\bar{y}_{\cdot ij\cdot}+\bar{y}_{t\cdots}+\bar{y}_{\cdot i\cdot\cdot}+\bar{y}_{\cdot\cdot j\cdot}-\bar{y}_{\cdots})^2$
Het. bet. reg.'s	$n-1$	M_ϕ	$q\sum_{i=1}^{n}[\sum_{t=1}^{m}\sum_{j=1}^{p}(\bar{y}_{tij\cdot}-\bar{y}_{ti\cdot\cdot}-\bar{y}_{t\cdot j\cdot}-\bar{y}_{\cdot ij\cdot}+\bar{y}_{t\cdots}+\bar{y}_{\cdot i\cdot\cdot}+\bar{y}_{\cdot\cdot j\cdot}-\bar{y}_{\cdots})(\bar{y}_{t\cdot j\cdot}-\bar{y}_{t\cdots}-\bar{y}_{\cdot\cdot j\cdot}+\bar{y}_{\cdots})]^2/\sum_{t=1}^{m}\sum_{j=1}^{p}(\bar{y}_{t\cdot j\cdot}-\bar{y}_{t\cdots}-\bar{y}_{\cdot\cdot j\cdot}+\bar{y}_{\cdots})^2$
Residual	$(n-1)(mp-m-p)$	M_θ	$SS_\tau-SS_\phi$
Error	$mnp(q-1)$	M_e	$\sum_{t=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q}(y_{tijk}-\bar{y}_{tij\cdot})^2$
Total	$mnpq-1$		$\sum_{t=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q}(y_{tijk}-\bar{y}_{\cdots})^2$

both linear and non-linear components of interaction variation are present in the data. In this case, the practical usefulness of any predictions will depend on the relative magnitudes of the two M.S.'s. This measure referred to as the "linear proportion" is proposed by Fripp and Caten (1971). It is defined as $[l/(l+nl)]\times 100$, where l is the heterogeneity between regressions M.S. minus the error mean square (linear component) and nl is the residual M.S. minus the error mean square

(nonlinear component). The heterogeneity between regressions M.S. is greater than its residual, indicating that a major part of interaction variation is accounted for by differences in these linear regressions. The reliable predictions can still be made.

In a two-way classification, the regressions approach of Lin and Thompson's (1975) method provides a natural and well-defined stopping criterion which would obtain groups with the same intercept (mean performance) and the same regression slope $(1+b)$. Since the dynamic model yields three regression coefficients for each genotype, the G × T, G × E, and G × T × E interactions should be separately considered in this cluster analysis. Dropping their criticism, let $\bar{y}_{.ij.}$ represents the mean observed performance of the i^{th} genotype in the j^{th} environment, then in a two-way classification, the original model (3) can be written as

$$\bar{y}_{.ij.} = \mu + G_i + (1+b_i)\,E_j + \delta_{ij} + \bar{e}_{.ij.} \tag{5}$$

and its reduced equation is

$$\bar{y}_{.ij.} = \mu_i + b'_i E_j + \Delta_{ij} \tag{6}$$

where $\mu_i = \mu + G_i$, $b'_i = 1 + b_i$, and Δ_{ij} is the residual from the linear regression of $\bar{y}_{.ij.}$ on E_j. It may be noted that if the variance of e_{ijk} is assumed homogeneous, the variance of the Δ_{ij} will also be homogeneous within genotypes but will be heterogeneous among genotypes if the b'_i were unequal. The dissimilarity index for a subset of r $(r \leqq n)$ genotypes $V_1, V_2, \ldots, V_r$ is defined by

$$d(V_1 V_2 \ldots V_r) = \left[SS(V_1 V_2 \ldots V_r) - \sum_{i=1}^{r} SS(V_i)\right] \Big/ 2(r-1)\,s^2 \tag{7}$$

where

$$s^2 = \sum_{i=1}^{n} SS(V_i)/[(n-1)(p-2)],$$

and $SS(V_1 V_2 \ldots)$ represent the sums of squares of residuals for genotype $V_1 V_2 \ldots$ Then, for example, $SS(V_i)$ simply represents the sum of squares of residuals for V_i. The statistic $d()$ is the variance ratio for testing the hypothesis of a common regression line against the alternative hypothesis of r independent regression lines. By definition,

$$d(V_1 V_2 \ldots V_r) = [2/r(r-1)] \sum_{i>i'}^{r} d(V_i V_{i'}), \tag{8}$$

that is, the dissimilarity index for r genotypes is equal to the mean of the indices for all C_2^r possible pairs of genotypes. It was shown by Lin and Thompson (1975) that

$$d(V_i V_{i'}) = \left[\sum_{j=1}^{p} E_j^2 (b_i' - b'_{i'})^2 + p(\mu_i - \mu_{i'})^2\right] \Big/ 4s^2 \tag{9}$$

where $\hat{\mu}_i = \bar{y}_{.i..}$. Hence, the grouping is equally based jointly on the M.S. of residuals from $\bar{y}_{.ij.}$ regressed on E_j, regression coefficient b'_i and mean performance μ_i.

If the regressions of α_{ti} on T_t are considered in a two-way classification, the above grouping procedure will be applied in a similar way, namely

$$\bar{y}_{ti..} = \mu_i + \xi_i' T_t + H_{ti} \tag{10}$$

where $\xi_i' = 1 + \xi_i$, and H_{ti} is the residual from the linear regression on T_t.

In the G×T×E interaction, there is a slight adjustment from above-mentioned. If the G × T × E interaction, $\bar{y}_{tij.}$, can be expressed as a linear function of the time-space index β_{tj}, we reduce the model (3) into

$$\bar{y}_{tij.} = \mu + V_i + T_t + E_j + \beta_{tj} + \phi_i \beta_{tj} + \theta_{tij} + \bar{e}_{tij.} \tag{11}$$

and obtain

$$\bar{y}_{tij.} - T_t - E_j = \mu_i + \phi'_i \beta_{tj} + \Theta_{tij} \tag{12}$$

where $\mu_i = \mu + V_i$, $\phi_i' = 1 + \phi_i$, and Θ_{tij} is the residual from the linear regression on β_{tj}. Since $\hat{T}_t = \bar{y}_{t...} - \bar{y}_{....}$ and $\hat{E}_j = \bar{y}_{..j.} - \bar{y}_{....}$, $(\bar{y}_{tij.} - T_t - E_j)$ can be estimated as $(\bar{y}_{tij.} - \bar{y}_{t...} - \bar{y}_{..j.} + 2\bar{y}_{....})$. By this way the method of Lin-Thompson will still be used to group the genotypes.

Finally, the smallest dissimilarity indices, min $d()$, constructed in each cluster cycle will be compared with the F value for r and $(n-1)(p-2)$ degrees of freedom at a predetermined probability level. Stopping the process when the min $d()$ exceeds the F value.

Materials and Methods

Twenty four genetical lines of *Arabidopsis thaliana* (L.) Heynh. (Cruciferae) were used in this study (Table 2). The experimental methods were the same as described in the previous paper (Wu, 1972). They were cultured with five replicates under twelve different environments.

The treatments consisted of six different artificial weather conditions as shown in Table 3. Each condition has two levels of fertilization, i.e., cultured in sand and added with water, with or without nutritions. Thus, the experiment involved a total of twelve environments.

The mean fresh weight of five plants from each line was recorded weekly at random from the first week to the seventh week, that is, from germination to the maturity of early flowering lines.

Table 2. *Experimental meterials*

Code No.	Race	Original or genetical description	Obtained from
1.	F 21	Chemical mutant	Dr. Robbelen
2.	C	X-ray mutant, flower after 22-25 days 6-7 rosette leaves, early flowering	Dr. van der Veen
3.	GR 1, 4	Langen 1 (Germany)	Dr. Oka
4.	SI-2	Catania (Sicily)	Dr. Oka
5.	AU	Graz (Austria)	Dr. Oka
6.	JA-2	Tsu (Japan)	Dr. Oka
7.	GR 2, 3	Langen 2 (Germany)	Dr. Oka
8.	Co-1	Coimbra (Portugal)	Dr. Robbelen
9.	Estland	Estland (Soviet Union)	Dr. Fujii
10.	Po-1	Poopelsdorf/Bown (Germany)	Dr. Oka
11.	EG-5	Early race of Langen 1	Dr. Oka
12.	En	Enkheim/Frankfurt (Germany)	Dr. Robbelen
13.	Wil-2	Wilna (Soviet Union), Zakret-park	Dr. Robbelen
14.	F 104	Chemical mutant	Dr. Robbelen
15.	LM-4	Late mutant from Landsberg	Dr. Oka
16.	51	EMS induced from C, 6-7 rosette leaves, late flowering than C	Dr. van der Veen
17.	Ch-1	Chlorophyll mutant	Dr. Robbelen
18.	φy-0	φystese/Hardanger Fjord (Norway)	Dr. Robbelen
19.	F 26	Chemical mutant	Dr. Robbelen
20.	LM-1	Late mutant from Landsberg	Dr. Oka
21.	51 A	Late mutant with large effect, i. e., flowering at least 10 days later and having at least 10 rosette leaves more than Race C	Dr. van der Veen
22.	51 D	Late mutant from Race 51	Dr. ven der Veen
23.	Hm	Hannoversh-monden	Dr. Hollamby
24.	A 136	Early flowering mutant of Estland	Dr. Hollamby

Table 3. *Artificial weather conditions*

	Light		Temperature (°C)	
	Intensity (lux)	Period (h)	Day	Night
1	10,000	14	25	20
2	24,000	14	27	22
3	5,500	14	30	25
4	natural light and natural day-length		25	20
5			20	20
6			30	25

Results

Analysis of Variance

The result of ANOVA is shown in Table 4. G × T, G × E, and G × T × E interactions were highly significant. This means that when genetical lines of *Arabidopsis thaliana* were grown in different environments throughout the whole growth period, the G × E interaction could be partitioned into: G × T, G × E, and G × T × E interactions. Further, similarly as found in the regression analyses for all interactions, the M.S. due to heterogeneity between regressions and the corresponding residual M.S. were both highly significant, indicating the presence of both linear and non-linear components of variation. However, the heterogeneity between regressions M.S. were greater than their respective residuals. The values of "linear proportion" were 96.79%, 79.61% and 93.11% for the G × T, G × E and G × T ×E interactions, respectively. The estimated values of the phenotypic mean and three kinds of regression coefficients, i.e. stability indices, for each genotype are presented in Table 5. Most of the regression coefficients significantly differed from zero (except for lines 4 and 22). Therefore, a major part of the interaction variation could be accounted for by the heterogeneity of the regressions of the individual genotypes.

Grouping of Genotypes Based on Linear Regression Approach

Since a high degree of linearity existed, the grouping procedure which combines

Table 4. *ANOVA (T, G, E are fixed model)*

Source	D.F.	M.S.	Linear proportion
Time (T)	6	530914.37**	
Enviroment (E)	11	52198.67**	
Genotype (G)	23	35760.45**	
T×E	66	7535.56**	
G×T	138	5050.98**	
Het. bet. reg.'s	23	24220.32**	96.79%
Residual	115	1217.12**	
G×E	253	3840.15**	
Het. bet. reg.'s	23	10969.17**	79.61%
Residual	230	3127.25**	
G×T×E	1518	1279.31**	
Het. bet. reg.'s	23	10102.25**	93.11%
Residual	1495	1143.58**	
Error	8064	427.29	
Total	10079		

** and *, significant at 1% and 5%, respectively.

Table 5. *The phenotypic mean ($\bar{y}_{\cdot i\cdot\cdot}$) and regression coefficients (ξ'_i, b'_i, ϕ'_i) for fresh weight of the 24 genetical lines in dynamic model*

Line(i)	$\bar{y}_{\cdot i\cdot\cdot}$	ξ'_i	b'_i	ϕ'_i
1	4.85	.17**	.21**	.10*
2	23.97	1.22**	1.49**	2.12**
3	25.92	1.09**	.49*	.45*
4	15.06	.59**	.22	.22
5	23.16	1.08**	.96**	.97**
6	24.68	1.10**	.41*	.49**
7	24.15	1.07**	.90**	.85**
8	12.96	.50**	.45**	.34**
9	22.87	.98**	1.07**	1.06**
10	37.57	1.55**	2.98**	2.29**
11	43.10	1.97**	2.30**	2.71**
12	20.20	.92**	.75**	.77**
13	22.74	.99**	1.40**	1.29**
14	18.35	.77**	1.11**	.89**
15	30.67	1.29**	.79*	.93*
16	18.22	.83**	.57**	.54**
17	15.50	.71**	.86**	.78**
18	31.24	1.46**	1.95**	2.21**
19	17.08	.80**	.81**	.86**
20	28.09	1.16**	.95**	.88**
21	20.98	1.00**	.87**	.98**
22	1.88	.05**	.01	.01
23	25.45	1.18**	1.23**	1.05**
24	32.37	1.53**	1.23**	1.20**
Mean	22.54			
LSD 5%	25.62			
LSD 1%	33.68			

**, *: significantly differ from zero at the 1% and 5% level, respectively.

the concept of testing for a common regression and the technique of cluster analysis was used.

For the G × T interactions, six groups were formed when the clustering was stopped after the smallest dissimilarity index exceeded the critical (5 per cent) F-value. These groups and their corresponding genotypes are illustrated in Table 6-(a). Grouping of genotypes would be helpful in selection of widely adapted varieties. As a result, genotypes within group III possessed the desirable

adaptability. They were characterized by a medium mean fresh weight of 23.84 ± 2.23 g, and a common regression coefficient of 1.07 ± 0.09. Group VI which included line 11 alone, exhibited yield superiority but showed low degree of adaptation.

Considering the G × E interactions, group III exhibited a distinct adaptation superiority over groups I, II and IV (Table 6-(b)).

The particular interest was the G × T × E interaction, whose pattern of variation among spaces are different over times, and the results showed that twenty-four genotypes were classified into five groups (Table 6-(c)). Group III with thirteen genotypes possessed satisfactory adaptability. They were charac terized by genotypes with the performance above the mean (25.02 ± 3.55 g), and linear regression near 1.00.

In all interactions, genotypes comprised three broad categories. The first category contained lines 1 and 22 with low means and above avarage stability. In the second category, which the genotypes had medium means and linear regressions near 1.00, i.e. general adaptability. The remaining groups formed the last

Table 6. *Grouping of genotypes by cluster method based on a linear function of the genotype-environment interaction*

Group	Coded No.	Min d( )	Mean	Regr. coef.
(a) For G×T interaction				
I	1 22	1.1560	3.37±2.10	.11±.08
II	4 8 14 16 17 19	1.6506	16.02±2.09	.70±.13
III	2 3 5 6 7 9 12 13 20 21 23	1.3159	23.84±2.23	1.07±.09
IV	15 18 24	.9499	31.43±.87	1.43±.12
V	10	2.3374	37.57	1.55
VI	11	7.5056	43.10	1.97
(b) For G×E interaction				
I	1 22	.3718	3.37±2.10	.11±.14
II	4 8 12 14 16 17 19 21	.7878	17.29±2.70	.71±.28
III	2 3 5 6 7 9 13 15 18 20 23 24	1.5000	26.28±3.46	1.07±.43
IV	10 11	1.9045	40.33±3.91	2.64±.48
(c) For G×T×E interaction				
I	1 22	.2192	2.37±2.10	.06±.06
II	4 8 14 16 17 19	.4724	16.20±2.09	.61±.28
III	2 3 5 6 7 9 12 13 15 20 21 23 24	1.1573	25.02±3.55	1.00±.41
IV	10 18	.9518	34.41±4.48	2.25±.06
V	11	2.5445	43.10	2.71

min d( ): the smallest dissimilarity index in the clustering cycle.
regr. coef.: ξ'_i, b'_i or ϕ'_i.

category within which included lines 10 and 11, the genotypes possessed high means and high regression coefficients (ξ', b' or ϕ') that were significantly larger than 1.00, indicating that they would be specially good under favorable environments but not under less favorable conditions. Hence, such a direct link of cluster method and linear regression approach seemed to be useful in the analysis of adaptation of *Arabidopsis thaliana* genotypes for testing fresh weight in this study.

Correlation Analysis

Table 6 illustrated that the phenotypic mean was positively associated with the three regression coefficients ξ', b' and ϕ', respectively. This could be detected by a correlation analysis as shown in Table 7. From results of Table 7, simple positive correlations between $\bar{y}_{.i..}$ and linear regressions, ξ'_i, b'_i, ϕ'_i were all significant. It seemed impossible to have high mean yield and above avarage stability in a genotype, i.e. an ideal genotype of Finlay & Wilkinson (1963). However, these associations were not absolute, particularly between $\bar{y}_{.i..}$ and b'_i($r = 0.7878$), and between $\bar{y}_{.i..}$ and $\phi_i{}'$($r=0.7985$), since there were a few genotypes with high performance in fresh weight and low linear sensitivity, and vice versa. This suggested that one can breed for genotypes with above mean performance and general adapatablity. Although there were significantly positive correlations occurred between the linear regressions from a simple correlation analysis, the high correlation between ξ'_i and b'_i disappeared when the phenotypic mean $\bar{y}_{.i..}$ was fixed.

Table 7. *Simple and partial (in parentheses) correlations among phenotypic mean $\bar{y}_{.i..}$ and regression coefficients of ξ'_i, b'_i and ϕ'_i*

	ξ'_i	b'_i	ϕ'_i
$\bar{y}_{.i..}$	.9880**	.7878*	.7985*
ξ'_i		.7850*(.0699)	.8296* (.4375*)
b'_i			.9418**(.8435**)

*, **: significant at 5% and 1% level, respectively.

Measurement of Phenotypic Stability

Breeding for stable varieties has received much attention recently. We would concentrated on stability rather than plasticity in this paper, even though *Arabidopsis thaliana* is a wild plant in which phenotypic plasticity plays a role in adaptation. The dynamic model provided three stability indices: the growth stability (ξ'), the environment stability (b') and the growth-environment stability (ϕ'). Since the growth stability was uncorrelated with the environment stability, the evaluation of G × T × E interaction of genotypes was equivanlent to simultaneous

consideration of G × T interaction and G × E interaction. The relationship among growth stability, environment stability and performance (fresh weight per plant) for each line is shown in Figure 1. It revealed a very striking configuration that the lines in the same group approximately the same response to time and space, and the resulting groups with above mean yield and satisfactory stability would be indicated by the cube in the center section on the left. This graphic summary could illustrate effectiveness of the above grouping in the dynamic model. The analysis of data in this study showed that no inbred lines could be selected when the criterion of selection was defined as high performance in fresh weight and perfect stability (indices ≤ 1.00). If the objective was to select individuals with the performance above the mean and general stability (indices near 1.00), the lines 3, 5, 6, 7, 9, 20, 23 could be acceptable. The results were consistent with the above grouping.

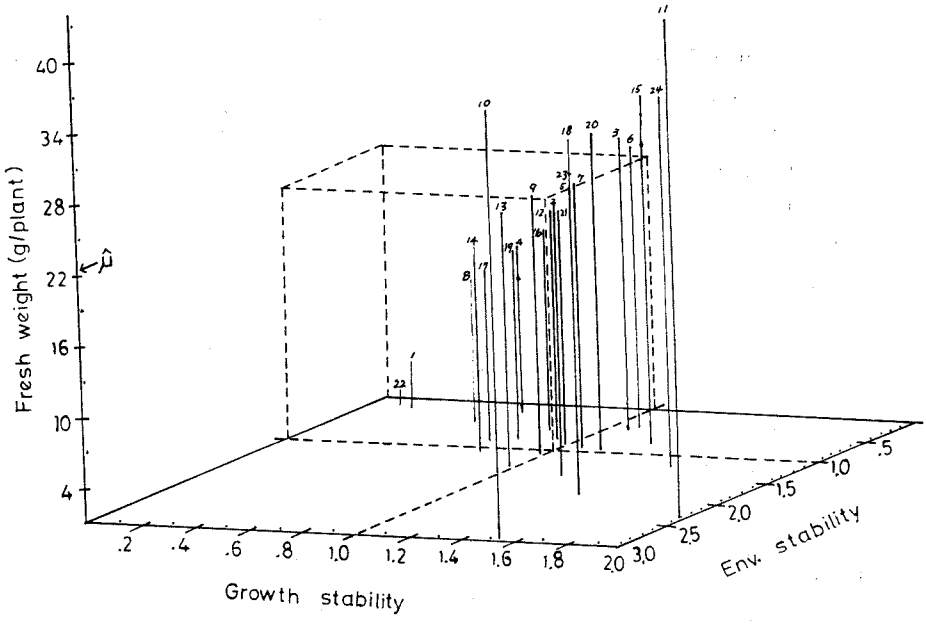

Fig. 1. The relationship among growth stability (ξ'_i), environment stability (b'_i) and performance ($\bar{y}_{\cdot i \cdot \cdot}$) of fresh weight for each line.

Discussion

Early workers used linear regression method to give a preliminary description of the genotype-environment interaction. It is customary to examine the plant response under different environments at a specific time of growing period, particularly at the harvest. From a practical point of view, it will obviously be valuable if plant response to a range of environments can be predicted from

existing genotype-environment regression data in the earlier stage of growth. To do this, Snoad and Arthur (1976) evaluated the feasibility of using Finlay and Wilkinson (1963) regressions as predictive tools for a number of characters in pea. But, failure has resulted from two main causes: either the data have not been sufficiently linear to permit the necessary degree of accuracy, or when good linear data have been obtained, the degrees of response have differed from one time to another time. Thus, in this study, imposing a time factor upon the linear model showed that all the interactions (including G × T, G × E and G × T × E interaction) could be accounted for by a linear function of the corresponding indices (T, E and T × E), respectively. This suggests that the response of each genotype accross environments for different growing period can be approximated by a straight line, and the non-parallelism of respone may be indicated by the difference among slopes. Such a dynamic linear model will undoubtedly be useful to reduced complex interactions to a series of orderly linear responses.

The statistic is useful in assessing the relative degree of adaptation among genotypes, but it provides no information on similarity of response. A direct link between a cluster method and the linear regression was then considered. Clearly, the grouping method can simplify the interpretation of the data by reducing the number of regression lines for genotypes to the number of groups. From the resulting groups, a further consequence is that the genotypes also comprised three broad categories as: (a) low yield and high stability, (b) medium or above-mean yield and general stability, (c) high yield but plastic. This makes it easier to identify the zones of similar genotypes adaptation, particularly if the number of genotypes is very large. However, where a low degree of linearity exists, the grouping may be misinformative. It should be noted that no claim can be made of the uniqueness of the groups derived, since a change in the numerical measure may produce a different classification. We therefore do not suggest that such grouping method should be the sole means for classification of genotypic response. Here making no attempt to offer a critique of some cluster methods, we are only concerned to demonstrated the application of the method and consider what general conclusions may be drawn. In general, if the group structure reflects sensibly existing biological and plant breeding knowledge of the population, the technique may be regarded as applicable in the study.

The positive correlation between the phenotypic mean and the regression coefficient has been found in a number of previous studies, such as *Nicotiana* (Perkins and Jinks, 1968), *Hordeum* (Paroda and Hayes, 1971), *Arabidopsis* (Westerman, 1971), *Dactylis* and *Festuca* (Wright, 1971), *Schizophyllum* (Connolly and Jinks, 1975), and potato cyst nematode (Phillips *et al.*, 1979). Wright (1971, 1976) pointed out that the correlation to occur is that the Finlay and Wilkinson (1963) regressions should be concurrent, and pass through a common point. If the

regression are concurrent, the increase in yield brought about by genetic improvement will necessarily entail a loss of yield adaptation. However, breeding experiments have shown that relationship between the mean performance and the linear sensitivity can be broken by appropriate selection procedures (Brumpton *et al.*, 1977; Jinks *et al.*, 1977). It therefore becomes necessary to expain why the regression lines for different genotypes should be concurrent. Faced with this problem, an interpretation has been given by Hardwick (1981): owing to canalization or homeostasis, in certain environment where selection pressures are strongest the differences in performance among genotypes disappear. In fact, the yield and the linear regression are often positively correlated. The reason is that the stability in the regression model is a relative measure depending on the genotypes included in the test, because the mean of all genotypes is used as standard response in each environment. A genotype stable by this definition is so only with respect to the other genotypes in the test, without any assurance that it will appear stable if assessed with another set of genotypes. That is, the regression model for G × E interaction is a descriptive model based on the data being analysed, but not a predictive model. If a researcher is interested in comparing relative stability among the group of cultivars included in the experiment and if the regression model fits the data, the linear regression coefficient, i.e. stability index, is useful.

The main advantage of using the present dynamic model is that both G × T interaction and G × E interaction can be studied simultaneously. The model provides three kinds of stability indices, namely growth stability index, environment stability index and growth × environment stability index. A criterion of selection can be determined according to the breeding objective and it provides a direct and easy method of screening genotypes in different time and space. Like every other model, it will sometimes arise theoretical argument. The assumptions underlying the analysis of variance are that the experimental errors are normally and independently distributed, with the same error variance. Since the character for measuring in the present study was fresh weight, the single plant sampling for statistical analysis should be different throughout the growth periods. Thus, for the dynamic model considered here, serial correlations of the resulting experimental errors might be disregard. However, the error mean squares would not be homogeneous in different stage of growth and development. The departures from assumptions underlying the analysis of variance affect both the significance level and power of statistical tests. But the effect is minor when the assumptions are not exactly satisfied. That is, slight departures from the assumptions are of little concern. The usual approach to dealing with this situation of heterogeneity of variances is to transform the observations and apply the analysis to the transformed data. But, one should note that although transformation is theoretically sound, breeders have used it infrequently. The main reason is that the conclusions

of the analysis for transformed data are often difficult to be interpreted from a biological concept. Such a subject of how to select an appropriate transformation or utilize other approach to minimize the heterogeneity of variances in the present material would well repay further research.

Literature Cited

Brumpton, R. J., H. Boughey, and J. L. Jinks. 1977. Joint selection for both extremes of mean performance and of sensitivity to a macroenvironmental variable. I. Family selection. Heredity **38**: 219-226.

Bucio Alanis, L. and J, Hill. 1966. Environmental and genotype-environmental components of variability. II. Heterozygotes. Heredity **21**: 399-405.

Connolly, V. and J. L. Jinks. 1975. The genetic architecture of general and specific environmental sensitivity. Heredity **35**: 249-259.

Eberhart, S. A. and W. A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci. **6**: 36-40.

Finlay, K. W. and G. N. Wilkinson. 1963. The analysis of adaptation in a plant breeding programme. Aust. J. Agric. Res. **14**: 742-754.

Fripp, Y. J. and C. E. Caten. 1971. Genotype-environmental interactions in *Schizophyllum Commune*. I. Analysis and character. Heredity **27**: 393-407.

Hardwick, R. C. 1981. The analysis of genotype × environment interactions: what does it mean if varietal stability is linearly related to varietal performance? Euphytica **30**: 217-221.

Jinks, J. L., N. E. M. Jayasekara, and H. Boughey. 1977. Joint selection for both extremes of mean performance and of sensitivity to a macro-environmetal variable. II. Single seed descent. Heredity **39**: 345-355.

Lin, C. S. and B. Thompson. 1975. An empirical method of grouping genotypes based on a linear function of the genotype-environment interaction. Heredity **34**: 255-263.

Lu, H. Y. and H. P. Wu. 1986. Studies on the genotype-environment interaction of *Arabidopsis thaliana*. Bot. Bull. Academia Sinica **27**: 187-207.

Paroda, R. S. and J. D. Hayes. 1971. An investigation of genotype-environment interactions for rate of ear emergence in spring barley. Heredity **26**: 157-175.

Perkins, J. M. and Jinks, J. L. 1968. Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. Heredity **23**: 339-356.

Phillips, M. S., J. M. S. Forrest, and A. M. Hayter. 1979. Genotype × environment interactions for resistance to the white potato cyst nematode (*Globodera pallida*, pathotype E) in *Solanum vernei* × *S. tuberosum* hybrids. Euphytica **28**: 515-519.

Snoad, B. and Arthur, A. E. 1976. The use of regression techniques for predicting the response of peas to environment. Theor. Appl. Genetics **47**: 9-19.

Westerman, J. M. 1971. Genotype-environment interaction and developmental regulation in *Arabidopsis thaliana*. II. Inbred lines; analysis. Heredity **26**: 93-106.

Wright, A. J. 1971. The analysis and prediction of some two factor interactions in grass breeding. J. Agric. Sci., Camb. **76**: 301-306.

Wright, A. J. 1976. The significance for breeding of linear regression analysis of genotype-environment interactions. Heredity **37**: 83-93.

Wu, H. P. 1972. Genetic basis of plant stability in *Arabidopsis thaliana*. I. Fluctuation of heritability and correlation. Bot. Bull. Academia Sinica **13**: 29-36.

Yates, F. and W. G. Cochran. 1938. The analysis of groups of experiments. J. Agric. Sci., Camb. **28**: 556-580.

植物數量性狀之適應性動態模式之研究

I. 時空介量之估計

呂秀英 鄔宏潘

中央研究院植物研究所

直線廻歸模式能有效地估計數量性狀對環境的適應性，而導出多種應用方法。但是植物的數量性狀，常因生長及發育的不同而隨時間改變，故其適應性亦因生長發生變動。因此本研究針對此問題，在直線模式中導入時間因素，誘導出測定適應性的動態模式，並以實驗植物 *Arabidopsis thaliana* 爲材料，檢討所建立的理論模式的適用性。

其結果顯示，該模式確適合用以闡釋植物性狀在不同生長階段的適應性變化趨勢。同時考慮適應性與相似性，配之以分群方法，將具有相同廻歸係數及表現型均值的基因型歸併成組，使能充份區別系統間之反應差異，而能利用於育種的選拔工作。本研究的動態模式中各基因型均含有三種廻歸係數（穩定性係數），其中生長穩定性係數與環境穩定性係數間互相獨立，而這些係數與表現型均值之間則呈顯著的正相關關係。