# 作物基因型與環境交感

呂 秀 英
台灣省農業試驗所農藝系

## 前 言

數量性狀之表現型值(P)，是由基因型值(G)、環境變異(E)，以及基因型與環境(G X E)交感效應累加而成，即：

$$P = G + E + G \times E$$

而檢驗 G×E 交感效應之存在性，可藉試驗設計(design experiments)之各種變方分析法，在逢機型假設前提下，利用期望均方 (expected mean squares) 以估算其變方大小。基因型之 G×E 交感效應小者，其對環境的適應能力較大。該適應性(adaptation) 中尙包括穩定性(stability)，競爭性(competition)，忍耐性(tolerance)，可塑性(plasticity)等性質。其定義之區分如下(鄔 1972)：

穩定性 — 指二種以上不同的基因型(品種)在二種以上不同的環境下生活時，因環境的不同，使不同基因型對環境變異的反應產生相對間的差異。

競爭性 — 二種以上的基因型，在同一有限的環境下共同生長時，各基因型間產生的干擾現象。

忍耐性 — 同一基因型在不同特殊環境生長時，在各環境下表現值變化的大小。

可塑性 — 某一基因型由原來環境移到極端相異的環境下生長某一時期後，表現型產生顯著變化後再到原來環境下繼續使其生長，而其恢復能力的大小。

這些現象全是基因型與環境間發生的交感現象。

就人類的經濟立場而言，作物品種改良目的在增進其經濟性狀的生產力，且所選育的高產量品種最好能適應於各種不同的地理環境，因此在育種研究上多著重於生產力的穩定性(stability)；因而，在探討作物基因型與環境交感效應時，亦多只限於討論穩定性的問題。

因表示穩定性的大小，係以 G×E 交感效應一項爲分析對象；自1938年迄今約五十年的時間，各方學者所提出的方法已不下數十種，依其發展時間的先後及著眼點之不同，大致可分爲以下五大類來加以說明(呂 1988)：

(一)直線迴歸分析法
(二)變方成份分析法
(三)分群方法
(四)多變量分析法
(五)非介量分析法

其中第(一)及第(二)類別內的統計量在基本上乃互爲等義，其所用以評估穩定品種所得結論大致相近；一般以直線迴歸分析法最爲常用而有效，而當直線式不存在時，則建議可採用與其相關的變方成份統計量，其中又以變異係數(CV, Francis and Kannenberg 1978)使用最爲頻繁。分群方法因偏重於衆多基因型或環境上的分類，故甚少單獨將之應用在基因型與環境交感之研究上，而多配合第(一)或第(二)類的統計量使用。至於近年來發展的多變量分析，由於計算繁雜且結果較不符合生物學上之解釋，雖有其應用上的場合，但實際引用於作物育種上者並不多，而多侷限在統計方法及理論方面的探討。 Nassar and Hühn (1987) 提出的非介量分析法，具有毋須考慮觀測值是否符合常態分布的優點，在基因型與環境數均甚少之情形下，爲評估基因型之穩定性提供了一個相當理想的方法，然其適用性及可靠性，猶待更多的試驗研究予以證實及肯定( 呂 1989)。

迄今，大多數育種工作仍以迴歸係數作爲選拔的依據。不過，使用迴歸分析法仍不免有其統計上的限制與實用上的缺失，如何配合實際試驗資料之特性來適當運用，或當迴歸分析法不再適用時，應如何尋求合理的解決方法，都是吾人在探討作物基因型與環境交感效應時，極須要關切的課題。因此，本文即針對此問題，特別就迴歸分析法在闡釋交感效應上的應用情形，作一概要而全盤性的介紹，並就其統計上之前提及在實際應用場合中常常會碰到的問題及缺失，予以詳盡地探討，而提供其可能可以解決的方法，俾使吾人在研究作物基因型與環境交感之育種工作上能有所參考。

## 直線迴歸模式的建立

1938 年 Yates and Cochran 將每個大麥品種的平均產量，隨某環境(區域、季節)下之平均產量作迴歸，發現二者間具有高度的直線關係。

繼之， Finlay and Wilkinson 於 1963 年提出了一種在參試基因型甚多的前提下，以各環境之平均收量與整個試驗的總平均收量之偏差，作爲該一試驗的環境效應估值，稱之爲環境指標( environment index)，而將各基因型收量隨這些指標作迴歸的分析， 並以最小平方法 (least squares method)估計迴歸係數($b_i$)，其迴歸係數平均值 ($\bar{b}$)爲 1； 利用這些迴歸係數和品種之平均表現可以評估基因型的穩定性及生產力。

Eberhart and Russell (1966) 認爲除了以迴歸係數 $b_i$代表穩定性外，應再求迴歸係數值之變異性(以迴歸剩餘機差項之均方 $\delta'^2_i$ 估計之)，如果變異性甚小，始可認定迴歸係數確足夠表示各基因型對環境之適應性。

Perkins and Jinks (1968a) 則在進行迴歸分析之前先對各基因型收量作一區域效應(location effects)的矯正，即以 G×E 交感成分隨環境指標的迴歸分析($\beta_i = b_i - 1$)來解釋之，其所有迴歸係數之平均值($\overline{\beta}$)爲0。

Tai (1971)認爲上述環境效應估值，乃由性狀觀測值估算而得，不滿足一般迴歸分析中自變數爲固定值的假設，必須在交感項及環境指標符合聯合常態分布 (joint normal distribution)之前提下，使用最大概度估計法(maximun likelihood method)估得最大概度估值$\hat{\alpha}_i$及$\hat{\lambda}_i$，$\alpha_i$測量基因型 i 環境效應的直線反應，及$\lambda_i$爲以機差變方爲單位，測量離直線反應之距離大小。Tai 氏定義這兩種統計值爲基因型穩定性介量(genotypic stability parameters)的估值。此法

一般被認爲是 Eberhart and Russell 之方法在環境指標爲逢機型時的特殊型態。

以上方法之擬定均建立在直線關係成立的基礎上。其中迴歸機差均方 $\delta^2_i$ 主要是在說明模式滿意的程度，當 $\delta^2_i$ 太大時，表示迴歸模式不適合用以估計穩定性，須考慮其他的分析方法，故與基因型的穩定性無直接關係，不適合作穩定性之估值。而因直線迴歸模式係用以描述(descriptive)存在之 G×E 交感效應，其所代表之穩定性爲一種相對性的概念，無法將所得結果作任何的預測(prediction)，故就一有效預測模式而言，須首先採用獨立的環境變數後，才有必要再以剩餘均方 $\delta^2_i$ 之大小來探明其預測的可靠性 (Lin *et al.* 1986; Westcott 1986)。

## 直線迴歸模式的應用

利用直線迴歸分析法來解釋穩定性的問題，是目前最常用而有效的方法。由統計上之觀點，其雖有一些前提之限制，但與其他方法比較，在實際應用上確具有操作簡單與符合生物解釋的優點。

使用迴歸分析法須在基因型及環境個數夠多的情形下，滿足兩個統計前提：

1.直線關係成立(即複迴歸決定係數 $R^2_i$ 值或直線比例值高)，

2.所有基因型之迴歸剩餘機差均方互爲均質(homogeneous)；

倘若直線式不存在，迴歸係數便不足以代表基因型對環境之適應性；而當均方異質(heterogeneous)時，因各迴歸係數估計值之可靠程度不一，進行迴歸係數間之差異性比較，亦殊不恰當。如這兩前提不能完全符合時，則建議可採用其他的統計方法(見下節說明)。

歸納起來，欲將直線迴歸分析法用以闡釋作物基因型與環境交感的一完整而正確的作法如下：

將 v 個基因型培育於 t 種不同的環境下，重覆 r 次試驗的資料，以下列數學模式表之：

$$Y_{ijk}=m+V_i+E_j+g_{ij}+e_{ijk} \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad (1)$$

其中，

$i = 1 \ldots\ldots\ldots v$

$j = 1 \ldots\ldots\ldots t$

$k = 1 \ldots\ldots\ldots r$

而，

$Y_{ijk}$：基因型(品種) i 在環境 j 中第 k 個重覆的觀測值。

m ：總平均，$\hat{m} = \sum_{i=1}^{v}\sum_{j=1}^{t}\sum_{k=1}^{r} Y_{ijk}/vtr = \bar{y}...$。

$V_i$ ：基因型 i 的遺傳效應，$\hat{V}_i=\bar{y}_{i..}-\bar{y}...$，而 $\bar{y}_{.i.}=\sum_{j=1}^{t}\sum_{k=1}^{r} Y_{ijk}/tr$ 。

$E_j$ ：環境 j 的效應，$\hat{E}_j=\bar{y}_{.j.}-\bar{y}...$，而 $\bar{y}_{.j.}=\sum_{i=1}^{v}\sum_{k=1}^{r} Y_{ijk}/vr$ 。

$g_{ij}$ ：基因型 i 與環境 j 之交感效應，$\hat{g}_{ij}=y_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}...$，而

$$\bar{y}_{ij.} = \sum_{k=1}^{r} Y_{ijk}/r$$ 。

$e_{ijk}$：試驗機差，而符合平均值爲 0，變方爲$\sigma^2$之常態分布。

並假定：

$\sum_{i=1}^{v} V_i = 0$，$\sum_{j=1}^{t} E_j = 0$，$\sum_{i=1}^{v} g_{ij} = \sum_{j=1}^{t} g_{ij} = 0$ ，亦即假定$V_i$，$E_j$ 及 $g_{ij}$均爲固定型效應 (Wright 1976)。

由於交感效應多呈環境指標的直線函數，

$$g_{ij} = \beta_i E_j \quad \cdots\cdots (2)$$

$b_i$爲迴歸係數$\beta_i$之估值，根據最小平方法，

$$b_i = \sum_{j=1}^{t} \hat{g}_{ij}\hat{E}_j / \sum_{j=1}^{t} \hat{E}_j^2 \quad \cdots\cdots (3)$$

而

$\sum_{t=1}^{v} b_i = 0$， $\delta_{ij}$ 爲剩餘機差 ，$\sum_{i=1}^{v} \delta_{ij} = \sum_{j=1}^{t} \delta_{ij} = 0$，

故直線迴歸模式可寫成

$$Y_{ijk} = m + V_i + E_j + \beta_i E_j + \delta_{ij} + e_{ijk} \quad \cdots\cdots (4)$$

(Perkins and Jinks 1968a)

$$= m + V_i + (1+\beta_i)E_j + \delta_{ij} + e_{ijk} \quad \cdots\cdots (5)$$

(Finlay and Wilkinson 1963)

$$= \mu_i + \beta_i' E_j + \delta'_{ijk} \quad \cdots\cdots (6)$$

(Eberhart and Russell 1966)

並且變方分析表如表一所示。

顯著性測驗爲：

$$F_{(1)} = MS_b / MS_\delta$$

$$F_{(2)} = MS_\delta / MS_e$$

(1) $F_{(1)}$及 $F_{(2)}$均不顯著：交感效應不存在。

(2) $F_{(1)}$顯著而 $F_{(2)}$不顯著：所有基因型的交感效應都可以用直線迴歸模式解釋之。

(3) $F_{(1)}$ 及 $F_{(2)}$同時顯著：有些基因型呈直線關係，有些則否，可能是因爲各基因型對環境反應不一致。

(4) $F_{(1)}$不顯著但 $F_{(2)}$ 顯著：交感效應對環境反應呈非直線關係。

當迴歸間差項與剩餘機差項之檢定同時顯著時 (即上述第(3)種情形)，則可利用 Fripp and Caten(1971)之直線比例(linear proportion)來概估其中呈直線之基因型所佔的比例，其定義爲

表一、變方分析表

| 變　因 | 自由度 | 平 方 和 | 均 方 | 均 方 期 望 值 |
|---|---|---|---|---|
| 環　境 $E$ | t-1 | $SS_E=vr\sum_{j=1}^{t}(\bar{y}_{.j.}-\bar{y}_{...})^2$ | $MS_E$ | $\sigma_e^2+\frac{vr}{t-1}\sum_{j=1}^{t}E_j^2$ |
| 基 因 型 $v$ | v-1 | $SS_v=tr\sum_{i=1}^{v}(\bar{y}_{i..}-\bar{y}_{...})^2$ | $MS_v$ | $\sigma_e^2+\frac{tr}{v-1}\sum_{i=1}^{v}V_i^2$ |
| 交感效應 $g$ | (v-1)(t-1) | $SS_g=r\sum_{i=1}^{v}\sum_{j=1}^{t}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{..})^2$ | $MS_g$ | $\sigma_e^2+\frac{r}{(v-1)(t-1)}\sum_{i=1}^{v}\sum_{j=1}^{t}g_{ij}^2$ |
| 迴　歸 $b$ | v-1 | $SS_b=r\sum_{i=1}^{v}[\sum_{j=1}^{t}(\bar{y}_{.j.}+\bar{y}_{...})(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...})]^2/\sum_{j=1}^{t}(\bar{y}_{.j.}+\bar{y}_{...})^2=r\sum_{i=1}^{v}b_i^2\sum_{j=1}^{t}E_j^2$ | $MS_b$ | $\sigma_e^2+\frac{r}{v-1}\sum_{i=1}^{v}\sigma\delta_i^2+\frac{r}{v-1}\sum_{i=1}^{v}b_i^2\sum_{j=1}^{t}E_j^2$ |
| 剩　餘 $\delta$ | (v-1)(t-2) | $SS_\delta=r\sum_{i=1}^{v}[\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...}-b_i(\bar{y}_{.j.}-\bar{y}_{...})]^2=SSg-SS_b$ | $MS_\delta$ | $\sigma_e^2+\frac{r}{v-1}\sum_{i=1}^{v}\sigma\delta_i^2$ |
| 試驗機差 $e$ | vt(r-1) | $SS_e=\sum_{i=1}^{v}\sum_{j=1}^{t}\sum_{k=1}^{r}(y_{ijk}-\bar{y}_{ij.})^2=SS_T-SS_E-SS_v-SS_g$ | $MS_e$ | $\sigma_e^2$ |
| 總　計 $T$ | vtr-1 | $SS_T=\sum_{i=1}^{v}\sum_{j=1}^{t}\sum_{k=1}^{r}(y_{ijk}-\bar{y}_{...})^2$ | | |

$$LP(\%) = [L/(L+NL)] \times 100 \qquad \text{.....(7)}$$

式中直線成份(linear component) $L = MS_b - MS_e$ ，

非直線成份( nonlinear component) $NL = MS_\delta - MS_e$ ，

若比例甚高，則所有基因型仍可適用直線迴歸分析法比較其迴歸係數大小；反之若比例不高，則須藉"機差相關法 (error correlation)"將基因型分組，再就各組分別討論之 (Perkins and Jinks 1968b; Lin and Thompson 1975; Lu and Wu 1986)，其方法如下：

任取基因型 i 及 i'，則剩餘機差($\delta ij$)間的相關係數設爲

$$r = \sum_{j=1}^{t}\delta_{ij}\delta_{i'j} \Big/ \left(\sum_{j=1}^{t}\delta_{ij}\sum_{j=1}^{t}\delta_{i'j}\right)^{1/2} \qquad \text{.....(8)}$$

組內基因型互爲正相關，$\delta_{ij}$同向，大小相近；組間則爲負相關，$\delta_{ij}$ 反向，大小相近；又$\delta_{ij}$大小若差異甚大，則互不相關。

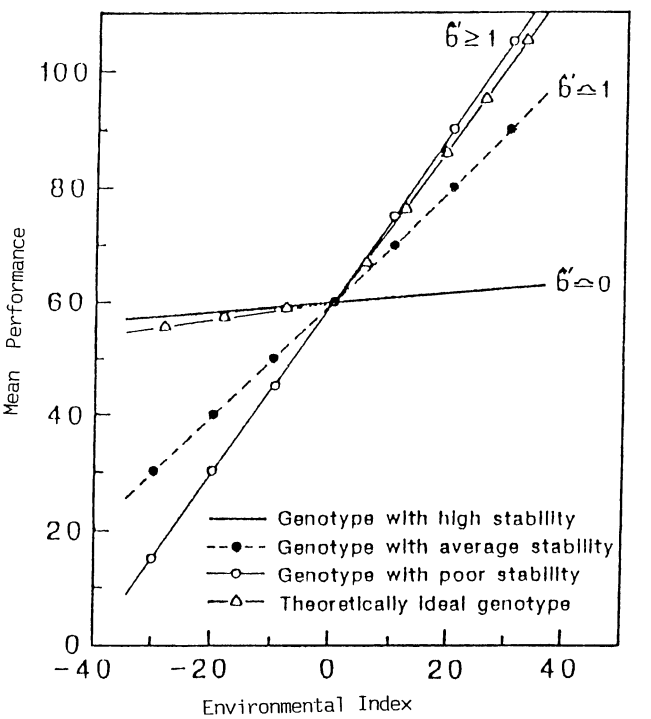
分組後，在進行以下變方分析，以檢定分組之效果，如表二所示。當組間與組內之均方比顯著時，表示分組有效。

表二、變方分析表

| 變　　因 | 自　　由　　度 | 組間均方 組內均方 V.R.[(i)/(ii)] |
|---|---|---|
| 基因型 | v-1 | |
| (i)組間 | g-1 | |
| (ii)組內 | v-g | |
| 迴歸間差 | v-1 | |
| (i)組間 | g-1 | |
| (ii)組內 | v-g | |
| 剩餘機差 | (v-1)(t-2) | |
| (i)組間 | g-1 | |
| (ii)組內 | (v-1)(t-2)-(g-1) | |

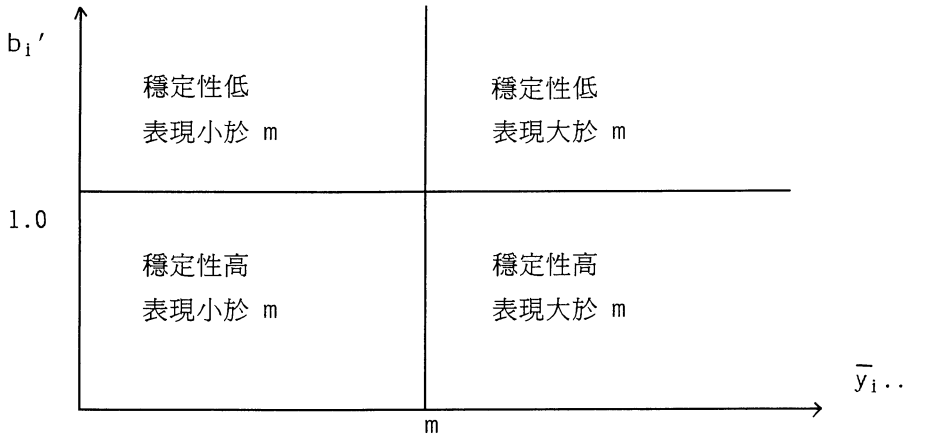
g : 組數　　V.R.: variance ratio

假設直線迴歸關係存在，利用 $b_i$或 $b_i'(=1+b_i)$ 作爲穩定性係數(兩者結果相同，而以後者表示者較多)以評價各基因型對環境的適應能力，如圖一所示。




圖一、 穩定性的類別 (Verma *et al.* 1978)

由於在育種研究上多著重於生產力的穩定性，因而，其實際應用上則配合平均表現值 (phenotypic mean)$\bar{y}_{i..}$ 來選拔穩定性高而表現良好的品種，如圖二所示。




圖二、 迴歸係數與平均表現值間之關係

## 直線迴歸模式使用上的限制及其解決方法

由於使用迴歸分析法必須在基因型及環境個數夠多的情形下，同時還要滿足兩個統計前提：(i) 直線關係成立，(ii) 所有基因型之迴歸剩餘機差均方互爲均質；但，並非所有試驗資料都能完全符合，因此在使用上仍須格外小心，否則計算過程雖然無誤，其評估結果也不一定正確。

一般常會碰到的問題及其可能解決方法如下：

(一)基因型及環境個數太少：

此易導致常態分布有所違離，剩餘機差項自由度太小等，以致迴歸係數之估算有偏差，而不足以代表基因型對環境之適應性。

最好的辦法是在試驗設計時，即應考慮清楚，盡量增加基因型及環境個數；一旦事前設計上未能周密，則在進行穩定性分析時已無法補救，而僅能以由平均值之變異或圖示來觀察參試基因型對參試環境的適應情形。

或者考慮非介量分析法，亦不失爲一理想的方法，但此法之適用性仍有待更多的試驗資料予以配合，始能獲得肯定。

(二)直線關係不成立：

建議採用變方成分統計量中最簡單的變異係數(coefficient of variability, CV)來評估各基因型之適應能力，

$$CVi = \frac{S_i}{\bar{X}_i} \times 100$$

，式中 $S_i$及 $\bar{X}_i$ 分別爲基因型i 之標準偏差及平均值。

至於其他非直線性的統計方法，只能用以描述某特定基因型對環境的變化趨勢，難以作不同基因型間之比較，或者計算過程較爲繁瑣，可能不利於實際運用。

(三)所有基因型之迴歸剩餘機差均方互爲異質：

事實上，該理論基礎係與 Eberhart and Russell之要求用迴歸剩餘機差項之均方 $\delta'^2_i$ 來評估迴歸係數值之變異性大小的主張，有相呼應之處， $\delta'^2_i$ 小(接近於 0)，則估得之迴歸係數才可靠，換言之，此時即滿足迴歸剩餘機差均方互爲均質(變方相等)之前提；但根據經驗，各基因型之迴歸剩餘機差均方能小至近於 0 的情形並不常發生 ，故原則上以剩餘機差均方須互爲均質的說法較爲合理。

當均方異質時，因各迴歸係數估計值之可靠程度不一，進行迴歸係數間之差異性比較，並不恰當。

但，這個前提多爲一般人所忽略。當均方爲異質時，統計上的解決方法是將資料予以適當轉換；不過轉換結果在生物學上往往很難解釋，而且資料轉換與否對穩定品種之選拔似乎影響也不大(呂及鄔 1989)，因此對該前提的重視程度猶待進一步地考榷研究。

## 參 考 文 獻


1. 呂秀英 1988 穩定性分析 科學農業 36:333-339.
2. 呂秀英 1989 非介量穩定性分析法之研究 I. 非介量與介量穩定性測量值間的比較 中華農業研究 38:208-215.
3. 鄔宏潘 1972 植物的適應性及其評價方法 科學農業 20:108-136.
4. 呂秀英、鄔宏潘 資料轉換對穩定性動態模式之影響研究 I. 時空介量之探討 中華農學會報 : 排版印刷中.
5. Eberhart, S. A. and W. A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci. 6: 36-40.
6. Finlay, K. W. and G. N. Wilkinson. 1963. The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. 14: 742-754.
7. Fripp, Yvonne J. and C. E. Caten. 1971. Genotype-environmental interactions in *Schizophyllum commune*. I. Analysis and character. Heredity 27: 393-407.
8. 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.
9. Lin, C. S. ,M. R. Binns, and L. P. Lefkovitch. 1986. Stability analysis: Where do we stand? Crop Sci. 26: 894-900.
10. Lu, H. Y. and H. P. Wu. 1986. Studies on the genotype-environment interaction of *Arabidopsis thaliana*. Bot. Bull. Academia Sinica 27 : 187-207.
11. Nassar, R. and M. Hühn. 1987. Studies on estimation of phenotypic stability : Tests of significance for nonparametric measures of phenotypic stability. Biometrics 43: 45-53.
12. Perkins, J. M. and J. L. Jinks. 1968a. Environmental and genotype-environmental

components of variability. III. Multiple lines and crosses. Heredity 23: 339-356.

13. Perkins, J. M. and J. L. Jinks. 1968b. Environmental and genotype-environmental components of variability. IV. Non-linear interactions for multiple inbred lines Heredity 23: 525-535.
14. Tai, G. C. C. 1971. Genotypic stability analysis and lts' application to potato regional trials. Crop Sci. 11: 184-190.
15. Verma, M. M., G. S. Chahal and B. R. Murty. 1978. Limitations of conventional regression analysis: a proposed modification. Theor. Appl. Genet. 53:89-91.
16. Westcott, B. 1986. Some methods of analysing genotype-environment interaction. Heredity 56: 243-253.
17. Wright, A. J. 1976. Bias in the estimation of regression coefficients in the analysis of genotype-environmental interaction. Heredity 37: 299-303.
18. Yates, F. and W. G. Cochran. 1938. The analysis of groups of experiments. J. Agric. Sci. 28: 556-580.